

# Description-Based Visualisation of Ethnic Facial Types

Sawitree Wisetchat

B.A., B.A., B.A., M.Phil.

Submitted in fulfilment of the requirements for the  
Degree of Doctor of Philosophy

School of Simulation and Visualisation  
The Glasgow School of Art

September, 2018

**SCHOOL OF  
SIMULATION AND  
VISUALISATION**  
**THE GLASGOW  
SCHOOL OF ART**



## Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification. No part of this work has been published.

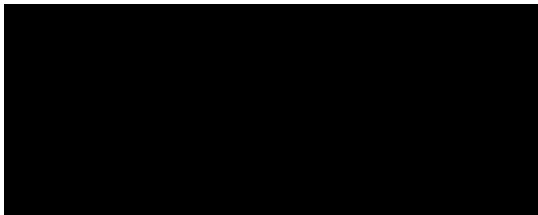
Signed,



Sawitree Wisetchat

Ph.D. Candidate

Date: 18/09/2018



Daniel Livingstone

Primary Supervisor

Date: 26/9/2018

## Abstract

This study reports on the design and evaluation of a tool to assist in the *description* and *visualisation* of the human face and variations in facial shape and proportions characteristic of different ethnicities. A comprehensive set of local shape features (sulci, folds, prominences, slopes, fossae, etc.) which constitute a visually-discernible ‘vocabulary’ for facial description. Each such feature has one or more continuous-valued attributes, some of which are dimensional and correspond directly to conventional anthropometric distance measurements between facial landmarks, while other attributes capture the shape or topography of that given feature. These attributes, distributed over six facial regions (eyes, nose, etc.), control a morphable model of facial shape that can approximate individual faces as well as the averaged faces of various ethnotypes. Clues to ethnic origin are often more effectively conveyed by *shape* attributes than through differences in anthropometric measurements due to large individual differences in facial dimensions within each ethnicity. Individual faces of representative ethnicities (European, East Asian, etc.) can then be modelled to establish the range of variation of the attributes (each represented by a corresponding three-dimensional ‘basis shape’). These attributes are designed to be quasi-orthogonal, in that the model can assume attribute values in arbitrary combination with minimal undesired interaction. They thus can serve as the basis of a set of dimensions or degrees of freedom.

The space of variation in facial shape defines an *ethnicity face space* (EFS), suitable for the human appreciation of facial variation across ethnicities, in contrast to a conventional *identity face space* (IFS) intended for automated detection of individual faces out of a sample set of faces from a single, homogeneous population. The dimensions comprising an IFS are based on holistic measurements and are usually not interpretable in terms of local facial dimensions or shape (i.e., they are not ‘semantic’). In contrast, for an EFS to facilitate our understanding of ethnic variation across faces (as opposed to ethnicity recognition) the underlying dimensions should correspond to visibly-discernible attributes. A shift from quantitative landmark-based anthropometric comparisons to local *shape* comparisons is demonstrated. Ethnic variation can be visually appreciated by observing

the changes in a model through animation. These changes can be tracked at different levels of complexity: across the whole face, by selected facial region, by isolated feature, and by isolated attribute of a given feature. This study demonstrates that an intuitive feature set, derived by artistically-informed visual observation, can provide a workable descriptive basis. While neither mathematically-complete nor strictly orthogonal, the feature space permits close surface fits between the morphable model and face scan data. This study is intended for the human visual appreciation of facial shape, the characteristics of differing ethnicities, and the quantification of those differences. It presumes a basic understanding of the standard practices in digital facial animation.

## Acknowledgements

I thank my supervisors Dr. Daniel Livingstone, Ms. Gillian Moffat, and Dr. Lisa DeBruine for their research guidance. I am grateful to Drs. Laura Gonzalez, Steve Love, and Nicky Bird at the Glasgow School of Art for their support and suggestions. Dr. Iris Holzleitner, Face Research Lab, Institute of Neuroscience and Psychology, University of Glasgow, provided valuable assistance with photogrammetric subject data. Drs. Scott Ernst and Kent A. Stevens, University of Oregon, provided technical assistance in Maya and scripting. I thank my examiners Dr. Nicky Bird and Dr. Robin Sloan, Abertay University, and Dr. Lynn-Sayers McHattie, convener, for their comments. I thank my friends and family for their patience and support, especially I am grateful to my mother.

All photographs by the author except as noted.

## List of Tables

Table 2.1. The anatomically-defined facial landmarks shown in Figure 2.	21
Table 2.2. A selection of absolute anthropometric measurements, in mm, derived from (Farkas et al., 2005). Males from two broad ethnicities (European and African) and three Asian ethnicities (Indian, Thai, and Japanese) are compared. Each mean and standard deviation is based on 30 samples.	22
Table 2.3. Anthropometric indices computed from (Farkas et al., 2005), with standard deviations for each quotient computed by error propagation of the sample means (Taylor, 1997). Males are compared for European versus African, African versus Japan and Japan versus Thai. Some comparisons are provided across ethnicities by unpaired t test. NS indicates not significantly different.	26
Table 2.4. Styles A-D refer to the styles of the four cups in Figure 2.4, from left to right. Four features are described, each with multiple possible modifiers. Two styles may be compared by by examining their respective columns. Styles B and C share similar rims and bases, but differ in profile and handle, for example.	37
Table 2.5. A matrix revealing similarities and differences in facial features of Buddha statues from (Wisetchat, 2011, table 3.1).	39
Table 5.1 Features of the eye region, with abbreviations, and their associated attributes. See Figure 5.3 for locations.	82
Table 5.2. Features of the mouth region, with abbreviations, and their associated attributes. See Figure 5.7 for location.	86
Table 5.3. Features of the nose region, abbreviations, and associated attributes. See Figure 5.10.	89

Table 5.4. Features and attributes of the cranium, mid-face, and jaw. Refer to Figure 5.14.	93
Table 8.1. Correspondence between facial attributes implicated in a genetic study (Adhikari et al., 2016) and the attributes of the Ethnicity Modeller.	174
Table 8.2. Comparison of the 17 attributes of the nose region for the two individuals modelled in Figure 8.1, sorted by descending absolute value of the difference in their values.	176
Table A3.1. Repeatability study results for matching the model mesh in the EM to three-dimensional face data (see Section 6.5.3 and Section 7.1).	246
Table A3.2. Accuracy study of five anthropometric attributes, comparing measurements taken with the Ethnicity Modeller for EAS_M20 with the corresponding measurements from Farkas et al. (2005) for Singaporean Chinese (CN) and Japanese (JP) males.	249
Table A3.3. Corresponding accuracy study for EUR_M20 compared to the anthropometric measurements from Farkas et al. (2005) for North American male (NA) and German male (DE).	250
Table A3.4. Computation of the differences in attribute values for two ethnicities. The table is sorted by the absolute value of the difference. Attributes in bold are calibrated and correspond to anthropometry measurements. Attributes of the nose region are highlighted (see text).	251
Table A3.5. Comparison of mean of 10 EAS male individuals with the corresponding measurements from Farkas et al. (2005) for Singaporean Chinese (CN) and Japanese (JP) males.	254

Table A3.6. Comparison of mean of 10 EUR male individuals with the corresponding measurements North American male (NA) and German male (DE) in (Farkas et al., 2005).	254
---	-----

## List of Figures

- Figure 2.1. The locations of the anthropometric landmarks defined in Table 2. 21
- Figure 2.2. Image with superimposed delineation. The white outlines trace a total of 189 defined fiducial points, some of which correspond to classical landmarks, others used to assist in bounding regions such as the eyebrows for subsequent image averaging 30
- Figure 2.3. Photographs of 11 AFR females were delineated and averaged to create the image AFR\_F11 (centre image, upper left). The mean template (lower left) does not reflect the large individual differences that are apparent when the individual templates are superimposed with a common interpupillary distance. Photographs by L. DeBruine, Face Research Laboratory, University of Glasgow, and the author. 31
- Figure 2.4. These four styles of coffee cup have obvious shape differences that are apparent in a glance. (They will be termed, from left to right, Styles A-D.) Careful scrutiny is required, however, to find their specific differences, which involves shifting the eyes repeatedly among them, to explore their shapes in order to find places where they differ, and where they share common features. 37
- Figure 2.5. Typeface styles such as Gill Sans (left) and Futura (right), are conventionally compared directly by direct scrutiny, alternately shifting visual attention between adjacent examples. 38
- Figure 2.6. These Buddha statues came from different geographic regions and cultures, and demonstrate distinctive styles. From left to right are examples from Pala, India, Sri Lanka, Pagan, Lan Na, and Dvaravati. Distinctive properties can be tabulated — see Table 2.5. 39



Figure 2.7. Sketches of different ethnicities (from left to right): Turkic, Margid Amerind, Nordic, Pacific Amerind, Baltic, South Chinese, and Japanese by Joumana Medlej (2012).	40
Figure 4.1. What ethnicities can one distinguish among these individuals? Clearly A is Western European and F is East Asian. There is also a West Asian (WAS, specifically Saudi), another Western European (WEU), an East African (EAF, i.e. Kenyan), and a South Asian (SAS, i.e., Pakistani). Answer: A = WEU, B = SAS, C = WEU, D = WAS, E = EAF, F = EAS.	62
Figure 5.1. Averaged faces of four ethnicities (top to bottom: EUR, SEA, EAS, and EUR), for male (left column), female (middle column), and 50% combined sexes (right column). This study focusses on the male face across ethnotypes, recognising that there are also sex differences within each ethnotype. Photographs by L. DeBruine, Face Research Laboratory, University of Glasgow, and the author.	80
Figure 5.2. Closeups of image-averaged eyes of four ethnicities, for both sexes. AFR_F11, for instance, is the average of 11 female AFR.	81
Figure 5.3. Locations of eye features. See Table 5.1 for definitions and associated attributes.	82
Figure 5.4. Closeups of the eyes of individuals of four ethnicities (top to bottom: EAS, SEA, AFR, and EUR).	83
Figure 5.5. Closeups of the eye region of four individuals for each of four ethnicities (top to bottom: EAS, SEA, AFR, and EUR).	84
Figure 5.6. Closeups of image-averaged mouths of four ethnicities, for both sexes. EAS_M13, for instance, is the average of 13 male EAS.	85

Figure 5.7. Locations of mouth features. See Table 5.2 for definitions and associated attributes.	86
Figure 5.8. Closeups of the mouth region of four individuals for each of four ethnicities (top to bottom: AFR, EAS, SEA, and EUR).	87
Figure 5.9. The nose region of four ethnicities, for both sexes.	88
Figure 5.10. Locations of nose features. See Table 5.3 for definitions and associated attributes.	89
Figure 5.11. Variations in the nose profile in four individuals for each for four ethnotypes (top to bottom: EUR, AFR, SEA, and EAS). Not to scale; each profile extends from just above the nasion to just below the philtrum.	90
Figure 5.12. Averaged faces of four ethnicities. Scaled to equal distance from eye to mouth.	91
Figure 5.13. Averaged images of four ethnicities in side view, aligned for constant eye height, and scaled to constant distance from eye to mouth.	91
Figure 5.14. Locations of features of the cranium, midface, and jaw. Refer to Table 5.4.	93
Figure 6.1. The base mesh rendered shown as a wireframe (left), as a polygonal mesh (middle), and as a smooth smooth surface (right).	95
Figure 6.2. Six facial regions — the cranium, eyes, nose, mid-face, mouth, and jaw — comprise a useful segmentation as few facial features span their boundaries. The face mesh is rendered as polygons (left) and as a smooth surface (right). Vertices along the boundaries of the regions are shared across regions and will be influenced by attributes in the neighboring regions.	99

Figure 6.3. The yellow rectangles represent copies of the base (with identical mesh topology but an attribute-specific shape). The green ellipses on the left represent blendshape nodes, each computing an interpolation between the maximum (on the left of the blendshape node) and the minimum (on the right) of that attribute. Next, these deformed meshes, each representing the contribution of an attribute (e.g., a philtrum length of 34%), deform the base mesh in parallel. 105

Figure 6.4. Screenshots of the Shape Editor. On the left the `Combined_blend` shape node is shown as organized into six groups (`Cranium`, `Eye`, etc.). On the right, the `Cranium` group is expanded to show its component targets (`CRN_depth_min`, `CRN_height_min`, etc.). Below `Combined_blend` is the group `Blends` (shown collapsed on the right, and expanded on the left) which organizes all of the blendshape nodes (`ALA_contour_blend`, `ALA_drop_blend`, etc.). 106

Figure 6.5. The base mesh (upper left) plus 10 unsigned attributes of the eye region. Each basis shape represents the extreme of the corresponding attribute (i.e., the maximum deformation relative to the base). Refer to Figure 5.3. 109

Figure 6.6. The basis shapes for four signed attributes governing the position of the endocanthus and the exocanthus (`ENC_x`, `ENC_y`, `EXC_x`, and `EXC_y`). Each attribute requires a pair of basis shapes to permit interpolation within the normalized range -1.0 to 1.0. A value of 0.3 for the `ENC_x`, for example, would be represented by a 0.3 interpolation between the basis shapes `ENC_x_min` and `ENC_x_max`. Refer to Table 5.1 and Figure 5.3. See Figure 6.6 for the remaining three signed eye attributes. 110

Figure 6.7. The remaining signed eye attributes, each represented by a pair of basis shapes within which the attribute is interpolated. Refer to Table 5.1 and Figure 5.3. 111

Figure 6.8. Models showing a variety of eye types created by different combinations of the 17 attributes of the eye region.	112
Figure 6.9. The base plus the three unsigned mouth attributes. The <i>IL_convexity</i> and <i>SL_convexity</i> attributes describe the subtly thicker tissue surrounding the lower and upper lips respectively. <i>PHL_protrusion</i> is the increased prominence of the philtrum just below the columella. Refer to Table 5.2 and Figure 5.7.	113
Figure 6.10. Six of the 12 signed mouth attributes, each represented by a corresponding pair of basis shapes. The mouth features are the cheilion <i>CH</i> , the cupid's bow <i>CPB</i> , the inferior labium <i>IL</i> , and the inferior labial tuberosity <i>ILT</i> (refer to Table 5.2 and Figure 5.7).	113
Figure 6.11. Basis shapes for the other four mouth features: the inferior vermillion border <i>IVB</i> , the philtrum <i>PHL</i> , the superior labium <i>SL</i> , the superior labial tuberosity <i>SLT</i> , and the superior vermillion border <i>SVB</i> . Refer to Table 5.2 and Figure 5.7.	114
Figure 6.12. Eight examples of mouth shapes, some chosen to show extreme values for the attributes in order to demonstrate the range of resultant shapes possible.	115
Figure 6.13. The base plus the four unsigned nose attributes are shown: <i>ALA_drop</i> (the relatively low placement of the outmost margin of the ala), <i>COL_drop</i> (the extent to which the columella hangs below the nostrils), <i>COL_show</i> (the elevation of the nostrils to reveal the nasal septum in side view), and <i>TIP_drop</i> (the relative displacement in Y of the very apex of the nose). The signed attributes are shown in Figures 6.14 and Figure 6.15. Refer to Table 5.3 and Figure 5.10.	116

- Figure 6.14. Eight of the 13 signed attributes of the nose (*ALA\_contour*, *ALA\_width*, *COL\_width*, *DSM\_length*, *DSM\_protrusion*, *DSM\_width*, *RAD\_protrusion*, and *RAD\_width*) are each represented by their extremes. For example, the width of the nose varies between that represented by *ALA\_width\_min* and *ALA\_width\_max*. Refer to Figure 5.10. 117
- Figure 6.15. The basis shape pairs for the remaining five signed attributes of the nose: *SID\_slope*, *TIP\_inclination*, *TIP\_pointed*, *TIP\_protrusion*, and *TIP\_width*. See also Figure 6.13 and refer to Table 5.3 and Figure 5.10. 118
- Figure 6.16. Examples of combinations of nose attributes. The profile is governed primarily by protrusion attributes while the breadth of the nose and the slope of the sidewall are determined by attributes that are quite independent of the protrusion attributes. A very large space of possible combinations result, some clearly unlikely exaggerations, others quite representative of observed nose shapes. 119
- Figure 6.17. The base plus the two unsigned cranial attributes *FOR\_protrusion* and *SOR\_protrusion*. Refer to Table 5.4 and Figure 5.14. 122
- Figure 6.18. The eight signed cranial attributes *CRN\_depth*, *CRN\_height*, *CRN\_width*, *FOR\_curvature*, *FOR\_slope*, *FOR\_width*, *SOR\_height*, and *TMP\_width*. Refer to Table 5.4 and Figure 5.16. 123
- Figure 6.19. Various combinations of the cranium attributes (see Table 5.4 and Figure 5.14). These eight examples are chosen to demonstrate extremes of combinations of these attributes. 124

- Figure 6.20. The seven signed attributes of the mid-face region. The cheek (*CHK\_protrusion* and *CHK\_width*) and the zygion (*ZYG\_protrusion* and *ZYG\_width*) have substantial spatial overlap, permitting a variety of smooth facial shapes in that region. The protrusion of the maxilla (*MAX\_protrusion*) governs the overall placement of the nose, mouth, and jaw regions. 125
- Figure 6.21. The base plus the three unsigned attributes of the jaw region: the weight of the labial mental sulcus (*LMS\_weight*) and the protrusion and width of the neck (*NCK\_protrusion* and *NCK\_width*). 126
- Figure 6.22. The seven signed attributes of the jaw region. This region has the following features: the chin or gnathion (*GNA*), the angle of the jaw or gonion (*GON*), and the mandible (*MDB*), with attributes governing the shape of the chin (*GNA\_x*, *GNA\_y*, and *GNA\_z*), the placement of the ‘angle’ of the jaw (*GON\_x* and *GON\_y*), and the protrusion and width of the mandible (*MDB\_protrusion* and *MDB\_width*). 127
- Figure 6.23. Examples of various combinations of jaw and midface attribute values. 128
- Figure 6.24. Modelling EAS\_M20, a data mesh representing the average of 20 male EAS. The closeness of fit is measured by 2 mm surface normals (yellow). On the left the normals are associated with the data mesh. On the right the normals are associated with the model mesh. By inspection the model mesh can be seen to fit generally well within  $\pm 2$  mm of the data mesh. 134
- Figure 6.25. Stereo-photogrammetric meshes for the computed average of 20 male EAS (top left) and 20 male EUR (bottom left), their corresponding models (center) and the two superimposed (right) to reveal their close similarity. 135

- Figure 7.1. As a ‘stress test’ on attribute orthogonality, the models from left to right show all 77 attributes set to their minimum values, all to their maximum values, then alternating minimum then maximum, and alternating maximum then minimum. The two faces on the left are maximally separated in this face space (they lie in diagonally opposite ‘corners’ of a 77-dimensional cube) as are the pair on the right. 147
- Figure 7.2. The results of matching the *Ethnicity Modeller* to stereo-photogrammetric scans of 10 individual EAS males. 157
- Figure 7.3. The results of matching the *Ethnicity Modeller* to stereo-photogrammetric scans of 10 individual EUR males. 158
- Figure 7.4. A sequence of blends that interpolate between two averaged images, representing EAS (left) and EUR (right) males, with intermediate values of 0.25, 0.5, and 0.75. See also (Wisetchat et al., 2018). 160
- Figure 7.5. The top row shows a blend from an AFR model to EUR\_M20, and the bottom row shows that AFR model morphing into EAS\_M20. In both cases the middle image is a 0.5 blend. In each case the images reflect interpolations from  $i = 0.0$  (left) to  $i = 1.0$  (right). See also ([Wisetchat, 2018](#)). 161
- Figure 7.6. Creating a caricature by extrapolation. The three images in the middle column show the  $i = 0.5$  interpolation between EAS and EUR from different perspectives. The columns to either side are the EAS ( $i = 0.0$ ) and EUR ( $i = 1.0$ ) models, and the extreme left and right columns show an exaggeration of EAS (left,  $i = -0.5$ ) and of EUR (right,  $i = 1.5$ ) relative to the mean ( $i = 0.0$ ). 162
- Figure 7.7. Progressive extrapolation. The facial characteristics of EAS (top left) and EUR (bottom left) are subject to 100% exaggeration (middle) and to 200% exaggeration (right). See also ([Wisetchat, 2018](#)). 163

Figure 7.8. The model of an average of 20 male EAS (EAS\_M20) on the left. The 165  
same model on the right now has EUR rather than EAS eyes (from EUR\_M20).

Figure 7.9. The model of an average of 20 male EUR (EUR\_M20) on the left, and 165  
the same model but with EAS eyes substituted, on the right. The eye attributes are  
from the average of 20 EAS (EAS\_M20, see Figure 7.8 left).

Figure 8.1. Two individuals that were modeled within the EM were then stored as 175  
very small files (each a dictionary of 77 attribute-value pairs). Since all models  
share the same mesh geometry and the same framework of blendshapes, a model  
constitutes a very economical encoding of a given face.

Figure 8.2. The middle image shows a 0.5 interpolation between a generic AFR 178  
and a EUR average EUR\_M20. (The specific AFR is a different individual than in  
Figure 7.5.). The models on the left and right have the same 50% AFR-EUR blend  
except on the left the mouth is 100% AFR and on the right the mouth is 100%  
EUR.

Figure 9.1. The average of 20 EAS males (left) and that of 20 EUR males (right) 187  
differ significantly in many facial attributes. In side-by-side presentation reveals  
the major differences become apparent by shifting attention between the two. A  
tabulation of the attribute values for the two models (Table A3.4) reveals more  
subtle differences that are not so readily noticed by comparing static images.  
Perhaps the most effective means to appreciate their ethnic differences is through  
observing an animated interpolation between the two models ([Wisetchat, 2018](#)).

Figure 9.2. Two examples of caricatures created through the *EM* interface. The 191  
manual, interactive effort could be replaced by a purely procedural process for  
automated character creation.



Figure A1.1. The initial state of the Ethnicity Modeller user interface.	215
Figure A1.2. The attribute editor consists of six tab panes. This shows the attributes for the eye. and their current values, with buttons to reset to zero and revert to last saved.	216
Figure A1.3. A tool for interpolating between two models, in this case EAS_M20 and EUR_M20, the averages of 20 male individuals each of EAS and EUR ethnicity. The model can either be interpolated with all attributes blending linearly, or on a per region basis independently.	220

# Glossary of Terms

## Abbreviations

AFR	<i>African ethnotype.</i>
ALA	<i>Ala Nasi.</i> The lateral portion of the nose which partly encloses the nostril (Figure 5.10). Attributes: <i>ALA_contour</i> , <i>ALA_drop</i> , and <i>ALA_width</i> (Table 5.3).
AL	<i>Alare.</i> The most lateral point on each alar contour, or lateral edge of nostril (Figure 2.1).
CN	<i>Chinese ethnotype.</i>
CH	<i>Cheilion.</i> The lateralmost point of the labial commissure, or corner of the mouth (Figures 2.1, 5.7). Attribute: <i>CH_x</i> (Table 5.2).
CHK	<i>Cheek</i> (Figure 5.14). Attributes: <i>CHK_protrusion</i> and <i>CHK_width</i> (Table 5.4).
COL	<i>Columella.</i> The tissue that separates the nares; inferior margin of the nasal septum (Figure 5.10). Attributes: <i>COL_drop</i> , <i>COL_show</i> , and <i>COL_width</i> (Table 5.3).
CPB	<i>Cupid's Bow.</i> Contour of the upper lip below the philtrum resembling a double-curved bow (Figure 5.7). Attributes: <i>CPB_depth</i> and <i>CPB_width</i> (Table 5.2).
CRN	<i>Cranium.</i> Region of head extending above forehead and laterally to temples (Figure 5.4). Attributes: <i>CRN_depth</i> , <i>CRN_height</i> , and <i>CRN_width</i> (Table 5.4).
CSV	<i>Comma-separated values.</i> A non-proprietary universal format for storing spreadsheet-organised data.
DE	<i>German ethnotype.</i>
DSM	<i>Dorsum.</i> The dorsal ridge of the nose extending from radix to tip (Figure 5.10). Attributes: <i>DSM_length</i> , <i>DSM_protrusion</i> , and <i>DSM_width</i> (Table 5.3).
EAF	<i>East African ethnotype.</i>
EAS	<i>East Asian ethnotype.</i>
ECF	<i>Epicanthal Fold.</i> A skin fold of the upper eyelid covering the inner corner of the eye nose (Figure 5.3). Attribute: <i>ECF_weight</i> (Table 5.1).
EFS	<i>Ethnicity Face Space</i> (Section 2.4.2).
EM	<i>Ethnicity Modeller</i> (Section 6.5).

ENC	<i>Endocanthion</i> . The medial (inner) commissure or corner of the eye fissure (Figure 5.3). Attributes: <i>ENC_angle</i> , <i>ENC_x</i> , and <i>ENC_y</i> (Table 5.1).
EUR	<i>European ethnotype</i> .
EXC	<i>Exocanthion</i> . The lateral (outer) commissure or corner of the eye fissure (Figure 5.3). Attributes: <i>EXC_y</i> and <i>EXC_y</i> (Table 5.1).
FOR	<i>Forehead</i> . The (Figure 5.14). Attributes: <i>FOR_curvature</i> , <i>FOR_protrusion</i> , <i>FOR_slope</i> , and <i>FOR_width</i> (Table 5.4).
GNA	<i>Gnathion</i> . Point on midline of most anterior projection of the chin (Figure 5.14). Attributes: <i>GNA_x</i> , <i>GNA_y</i> , and <i>GNA_z</i> (Table 5.4).
GON	<i>Gonion</i> . The most lateral point on the mandibular angle (corner of the jaw) (Figure 5.14). Attributes: <i>GON_x</i> and <i>GON_y</i> (Table 5.4).
IFS	<i>Identity Face Space</i> (Section 2.4.1).
IPC	<i>Inferior Palpebral Convexity</i> (Figure 5.3). Attributes <i>IPC_weight</i> (Table 5.1).
IPD	<i>Interpupillary Distance</i> (Figure 5.3). Attribute: <i>IPD_x</i> (Table 5.1).
IPS	<i>Infrapalpebral Sulcus</i> (Figure 5.3). Attribute: <i>IPS_weight</i> (Table 5.1).
IL	<i>Inferior Labium</i> . The lower lip distance (Figure 5.7). Attributes: <i>IL_convexity</i> , <i>IL_protrusion</i> , and <i>IL_thickness</i> (Table 5.2).
ILT	<i>Inferior Labial Tuberosity</i> (Figure 5.7). Attribute: <i>ILT_fullness</i> (Table 5.2).
IVB	<i>Inferior Vermillion Border</i> (Figure 5.7). Attribute: <i>IVB_curve</i> (Table 5.2).
JND	<i>Just noticeable difference</i> . The minimal perceptible adjustment to an experimental variable.
LMS	<i>Labial Mental Sulcus</i> (Figure 5.14). Attribute: <i>LMS_weight</i> (Table 5.4).
MAX	<i>Maxilla</i> . (Figure 5.14). Attributes: <i>MAX_protrusion</i> (Table 5.4).
MDB	<i>Mandible</i> . (Figure 5.14). Attributes: <i>MDB_protrusion</i> and <i>MDB_width</i> (Table 5.4).
N	<i>Nasion</i> . Deepest point below the glabella on bridge of the nose (Figure 2.1).
NCK	<i>Neck</i> (Figure 5.14). Attributes: <i>NCK_protrusion</i> and <i>NCK_width</i> (Table 5.4).
PF	<i>Palpebral Fissure</i> is the elliptic space between the medial and lateral canthi of the two open lids (Figure 5.3). Attribute: <i>PF_inclination</i> and <i>PF_width</i> (Table 5.1).

PHL	<i>Philtrum</i> (Figure 5.7). Attributes <i>PHL_length</i> , <i>PHL_protrusion</i> , and <i>PHL_width</i> (Table 5.2).
PMS	<i>Palpebromalar Sulcus</i> (Figure 5.3). Attribute: <i>IPC_weight</i> (Table 5.1).
PNS	<i>Palpebronasal Sulcus</i> (Figure 5.3). Attribute: <i>PNS_depth</i> (Table 5.1).
PTF	<i>Palpebrotemporal Fossa</i> (Figure 5.3). Attribute: <i>IPC_weight</i> (Table 5.1).
JSON	JavaScript Objective Notation.
JP	<i>Japanese ethnicity</i> .
NA	<i>North American ethnicity</i> .
WEU	<i>Western European ethnicity</i> .
WAS	<i>West Asian ethnicity</i> .
RAD	<i>Radix</i> . The bridge of the nose in vicinity of nasion (Figure 5.10). Attributes: <i>RAD_protrusion</i> and <i>RAD_width</i> (Table 5.3).
SAS	<i>South Asian ethnicity</i> .
SEA	<i>South East Asian ethnicity</i> .
SID	<i>Sidewall</i> (Figure 5.10). The lateral slope of the nose adjacent to the dorsum (Figure 5.10). Attribute: <i>SID_slope</i> (Table 5.3).
SOR	<i>Supraorbital Ridge</i> (Figure 5.14). Attributes: <i>SOR_height</i> and <i>SOR_protrusion</i> (Table 5.4).
SL	<i>Superior Labium</i> (Figure 5.7). Attributes: <i>SL_convexity</i> , <i>SL_protrusion</i> , and <i>SL_thickness</i> .
SLT	<i>Superior Labial Tuberosity</i> (Figure 5.7). Attribute: <i>SLT_fullness</i> (Table 5.2).
SN	<i>Subnasion</i> . Midpoint of the base of the nose where the columella meets the philtrum (Figure 2.1).
SPC	<i>Suprapalpebral Convexity</i> (Figure 5.3). Attribute: <i>SPC_weight</i> (Table 5.1).
SPS	<i>Suprapalpebral Sulcus</i> (Figure 5.3). Attribute: <i>SPS_weight</i> (Table 5.1).
STF	<i>Supratarsal Fold</i> (Figure 5.3). Attribute: <i>STF_weight</i> (Table 5.1).
SVB	<i>Superior Vermillion Border</i> (Figure 5.7). Attribute: <i>SVB_curve</i> (Table 5.2).
TIP	<i>Nasal Tip</i> (Figure 5.10). Attributes: <i>TIP_drop</i> , <i>TIP_inclination</i> , <i>TIP_pointed</i> , <i>TIP_protrusion</i> , and <i>TIP_width</i> (Table 5.3).
TMP	<i>Temple</i> (Figure 5.14). Attribute: <i>TMP_width</i> (Table 5.4).

**TRG** *Tragion*. Point on the upper margin of each tragus of the ear (Figure 5.14).

Attributes: *TRG\_y* and *TRG\_z* (Table 5.4).

**ZYG** *Zygion*. The most lateral point of each zygomatic arch (Figure 2.1, 5.14).

Attributes: *ZYG\_protrusion* and *ZYG\_width* (Table 5.4).

## Terminology (Specific to this Study)

*Attribute-value pair*. The association or binding of a specific value to a given attribute (see model).

*Basis shape*. A polygonal mesh representing a possible surface shape or configuration.

Basis shapes can be used in weighted combination to create a blend of multiple basis shapes (e.g., by a blendshape deformer, wherein the basis shapes are frequently termed ‘target shapes’).

*Deformer*. A digital animation technique for displacing the coordinates of a set of vertices comprising a polygonal mesh, causing that shape to smoothly morph.

*Error propagation*. The explicit computation of the uncertainty of a mathematical operation based on the explicit uncertainty of its operands.

*Ethnotype*. A geographically-defined population with shared physical characteristics.

*Extrapolation*. Estimation of values past a given range, effectively linear interpolation with a fraction  $< 0.0$  or  $> 1.0$ . Used to create exaggerations (Section 7.2.3).

*Face space*. The set of face configurations specified by all combinations of values along a specified set of dimensions (attributes or axes of variation).

*Fiducial point*. A distinguished facial location (e.g., the corner of the mouth) which can be located on multiple faces. Used to establish a common mapping for image averaging, for example.

*Holistic measurement*. A measurement that involves the entirety of an object, as opposed to a localised feature or region. Automated face recognition algorithms are frequently based on holistic measurements.

*Landmark*. A distinguished location on the face, either defined by a soft-tissue feature (such as the corner of the mouth, the cheilion) or a superficial point defined by an osteological feature (such as the nasion, glabella, or gonion).

*Linear interpolation.* Estimation of an intermediate value between two values as a specified fraction (0.0-1.0) of the difference between those two values.

*Model.* A set of attribute-value pairs specifying a particular face configuration. A model can be stored in JSON format and, when read, used to specify the attributes of a polygonal mesh for visualisation (Section 3.2.2).

*Morphing.* Deformation of a two- or three-dimensional shape, frequently by linear interpolation of the positions of its component pixels or vertices.

*Navigation (of face space).* Visualisation of faces that lie along a continuous-path trajectory in a face space. Linear interpolation morphing between two faces would follow a straight line path in face space.

*Polygonal mesh.* An approximation of a smooth surface by a collection of vertices, edges, and faces.

*Quasi-orthogonal.* Two attributes that act as approximately independent, i.e., can occur in arbitrary combination with minimal interference.

*Semantic.* A property, attribute, or measurement is semantic if its value has human-understandable meaning. The components of a colour space are semantic, while those derived by Principal Components Analysis are not semantic in this regard.

*Stereo-photogrammetric data.* A dataset of the three-dimensional coordinates of a surface derived by analysis of multiple photographic images taken from different positions.

*Topographic feature.* A description of the localised shape of a surface based on discrete distinctions of its geometry, such as folds, creases, concavities and convexities.

# Table of Contents

Abstract	i
Acknowledgments	iii
List of Tables	iv
List of Figures	vii
Glossary of Terms	xvii
Chapter 1. Introduction	1
1.1 Motivation	2
1.2 Research Objectives	5
1.3 Synopsis of the Thesis	7
1.3.1 Literature Review	7
1.3.2 Methodology	8
1.3.3 Facial Description	8
1.3.4 Facial Features and their Attributes	9
1.3.5 Developing the Parametric Model	9
1.3.6 Ethnicity Modeller Evaluation and Applications	10
Chapter 2. Literature Review	11
2.1 Facial Types	11
2.1.1 Categorising Ethnotypes Geographically	12
2.1.2 Categorising Ethnotypes Genetically	13
2.1.3 The Risk of Stereotyping	16
2.1.4 The Concept of an Average Face	17
2.2 Measuring Faces	19
2.2.1 Facial Landmarks and Measurements	20
2.2.2 Anthropometric Indices and Facial Proportions	23
2.2.3 Measurement Methods	27
2.3 Describing Faces	34
2.3.1 Descriptions versus Depictions	34
2.3.2 Categorical Descriptions	35
2.3.3 Visualising Shape Differences	42

2.4 Face Spaces	45
2.4.1 A Face Space for Recognising Individuals	46
2.4.2 A Face Space for Describing and Comparing Faces	48
2.4.3 Navigating Face Space	50
Chapter 3. Methodology	52
3.1 Creating a Proof-of-Concept Implementation	53
3.2 Synopsis of the Methodology	54
3.2.1 Defining Facial Attributes	54
3.2.2 Modeller Development	55
3.2.3 Evaluation	57
3.2.4 Model Manipulation and Ethnicity Visualisation	59
3.3 Working Assumptions	59
Chapter 4. Facial Description	60
4.1 Local Features	60
4.2 Attribute-Value Pairs	63
4.3 Description by Analogy	65
4.4 Orthogonality	65
4.5 The Insufficiency of Measurements as Descriptors	66
4.6 Variance and Measurement Error	67
4.7 Descriptions of Continuous-Valued Shape Attributes	69
4.8 Arbitrary but Useful	70
4.9 Exemplar-Based Attributes	71
4.10 Visualising Attributes	73
4.11 Defining a Face Space	75
Chapter 5. Facial Features and their Attributes	79
5.1 Eye Attributes	79
5.2 Mouth Attributes	85
5.3 Nose Attributes	87
5.4 Cranium, Midface, and Jaw Attributes	90



Chapter 6. Developing the Ethnicity Modeller	95
6.1 Defining Facial Attributes by Extremes	96
6.2 Facial Regions	98
6.3 Implementing Facial Attributes by Basis Shapes	100
6.3.1 Meshes and Deformers	102
6.3.2 Attribute Independence and Combinations	107
6.3.3 The Model Origin	108
6.4 Basis Shapes for Facial Attributes	108
6.4.1 Eye Attribute Basis Shapes	109
6.4.2 Mouth Attribute Basis Shapes	112
6.4.3 Nose Attribute Basis Shapes	116
6.4.4 Overlapping Features	119
6.4.5 Cranium, Midface, and Jaw Attribute Basis Shapes	121
6.5 The Ethnicity Modeller	129
6.5.1 The Model Data	130
6.5.2 Matching the Model to Two-Dimensional Images	131
6.5.3 Matching the Model to Three-Dimensional Data Meshes	132
Chapter 7. Ethnicity Modeller Evaluation and Applications	137
7.1 Evaluation	137
7.1.1 Measurement Precision	138
7.1.2 Measurement Accuracy	141
7.1.3 Attribute Orthogonality	144
7.1.4 Usability	146
7.2 Applications	154
7.2.1 Quantitative Comparison	154
7.2.2 Measurement of Individuals	156
7.2.3 Visualising Ethnic Differences	159
7.3 Summary of the Evaluation	166
Chapter 8. Discussion	169
8.1 About Face Space	170

8.1.1 An EFS Is Not an IFS	170
8.1.2 Attributes as EFS Dimensions	171
8.2 The Potential Genetic Basis for Facial Attributes	173
8.3 An Economical Facial Representation Scheme	175
8.4 Experimental Investigation of Ethnicity Recognition	177
8.5 Critical Evaluation	178
8.6 Summary	180
Chapter 9. Conclusions and Future Directions	182
9.1 Contributions	182
9.1.1 Regarding Ethnotypes and Face Spaces	182
9.1.2 Shapes and Basis Shapes	184
9.1.3 Visualising Ethnic Differences	185
9.1.4 Extending the Notion of Description by Example	188
9.2 Limitations	188
9.3 Future Directions	189
9.3.1 Creating Derivative Applications	189
9.3.2 Character and Caricature Design	191
9.3.3 Visualisation by Navigation	192
9.3.4 Other Domains	194
References	196
Appendix 1. User Interface	214
A1.1 Initialising the Ethnicity Modeller	214
A1.2 Model Editing	215
A1.2.1 Attribute Dependency	217
A1.2.2 Matching to Photogrammetric Data	218
A1.3 Ethnic Difference Visualisation	219
Appendix 2. Ethnicity Modeller Code	221
A2.1 The Configuration File	221
A2.2 The User Interface Launcher	222
A2.3 The Three-Dimensional Model	222

A2.4 The User Interface	228
A2.5 The CSV Exporter	243
Appendix 3. Quantitative Analyses	245
A3.1 Repeatability Study	245
A3.2 Accuracy Study	248
A3.3 Quantifying Ethnic Differences	251
A3.4 Measuring Individuals	253

# 1. Introduction

The human face exhibits both local and global variation across the globe. While the faces of people in any region vary from individual to individual, their faces share common features which vary gradually from one geographic region to another, such that far-separated populations display significant facial differences. The commonly-shared distinctive characteristics will be regarded as defining facial ethnotypes. This study focusses upon visualising and describing such ethnic variations in the human face. The emphasis here is on ethnic variation in the shape of the face, not other qualities such as eye and skin pigmentation, hair properties, and skin textures. By “visualising and describing”, the implication is that the descriptive language will have a *visual* component. Since the focus of this research is on helping humans — rather than computers — to appreciate and understand variations in the form of the face, the language that will be used must carry meaning about shape in terms that we can visualise.

This study represents an artist’s exploration of the face, substituting digital technology for the traditional sketch pencil. This work is interdisciplinary (e.g., drawing from anthropometry, digital image averaging, facial animation, and user interface and database scripting), but does not require specialised expertise in any of those areas. As an artist, my intention is to create a new descriptive framework with which to capture the patterns of facial variation across ethnicities, in a way that is intuitive yet specific and concrete. While holistic approaches have gained favour for recognising faces (Section 2.4.1), descriptions are not holistic. Instead, a face is naturally described in terms of its component parts, which is a task better suited for an artist than a mathematician. Artists are skilled in seeing the face as composed of a combination of folds, facets, contours and other local facial features, present in various combination in different ethnicities. The choice of facial features used in this study derive primarily from human facial anatomy. Someone with knowledge of human facial anatomy will thus find the terminology familiar. For others, the glossary and illustrations should be of assistance. A reader without basic knowledge

of the foundations of digital animation (such as polygonal meshes, linear interpolation, blendshape deformers) may find a few technical sections in Chapter 6 obscure in the details, but the gist should be clear nonetheless. Another reader with suitable technical background will readily understand how abstract facial attribute can be concretely represented through three-dimensional modelling and digital animation.

A proof-of-concept demonstration of the proposed scheme was created and evaluated to measure the precision, accuracy of the attribute representation scheme and the utility of the interactive user interface that was provided for manipulation of the attributes of the three-dimensional face model. This demonstration system was created only to evaluate the underlying concepts (representing facial attributes in terms of interpolation between extrema, achieving orthogonality across attributes, creating a novel, intuitive face space, etc.). While the details of the implementation are provided in an appendix for purposes of explanation and replication, this study is conceptual, not immediately practical (i.e., no stand-alone ‘app’ is offered). Knowledge of digital animation, Autodesk Maya, and Python would be required to fully understand these aspects of the thesis.

## **1.1 Motivation**

Faces are very important visually, as they tell us who someone is, what they are feeling, what they are thinking, where they are from, who they may be related to, their sex, their age, and so forth. We are expert at extracting such information in a glance. The same face which conveys social information through dynamic expressions and gestures continues to convey information when it is completely still, or photographed, or sketched, or modelled in clay. Even the simplest sketches or simplified solid sculptures or blurred images capture much of this information. A depiction of a face allows one to exaggerate, or alternatively, to de-emphasise or ‘average-away’ information about individuality, expression, sex, and so forth. Much

of my personal motivation for this research comes from my fascination regarding the face and the information it conveys to us.

Virtually any depiction of the face seems to convey so much information that it is fascinating to study what features of the face are analysed as we observe it. Much is being learned about facial clues to masculinity, attractiveness, social communication, and new technology is rapidly advancing for automating the recognition and classification of individuals, ‘reading’ where people are looking, their facial expressions, and so forth. The *study* of the human face is also simply fascinating.

I have been attracted to research into the visualisation and description of faces from an artistic perspective, and from the perspective that comes from living in two very different cultures, the West and the East. While all faces share many common, universal properties, there are also clues to ethnic origin that can be ‘read’ despite all the variation from one individual face to another. Artists are sensitive to these clues and can capture them and convey them through three-dimensional sculptures and various two-dimensional media.

Artists, physical anthropologists, surgeons, forensic investigators, and many others can describe in words what they see in faces, but any written or verbal description of a face is only able to make a small amount of that information explicit. Words are mostly used to direct an observer’s attention to one or another aspect or feature, but not so much to fully describe that feature. When asked to describe anything fully or in detail, we may use a few technical or specific terms, but we often rely on demonstrations, or analogies, or examples to fill in for what is missing. Experts use specialised terms (such as the medical name for some facial feature), but when one tries to track down their meaning, that trail often ends in an illustration, or perhaps a pair of illustrations that define the term visually. This seems circular. To describe some visual feature, the experts use a specific technical terminology, but those more

primitive terms are themselves ultimately defined *visually*, i.e., with reference to illustrations or visual exemplars.

Earlier I studied the evolution of the Buddha statue (Wisetchat, 2011, 2013a, 2013b) using digital technology to visualise how stylistic features of the statue have changed as Buddha statues were first created in India, then spread to Sri Lanka and Burma, then northern and central Thailand, with influence from Khmer and other cultures. A technical lexicon of facial features has been developed by art historians, and very much like the use of technical terms in the medical description of the human face, these statues are analysed and described for clues to social expressions (messages conveyed by the Buddha images, which has been found to differ with different Buddhist cultures), and different aesthetics towards beauty, masculinity, and so forth — all very similar to how the living human face is studied and described. The experts rely on visualisation to fill in for their written descriptions. A book on Buddha statues would be difficult to write without accompanying images. The role of the artist becomes very important in noticing the salient features and in conveying them by illustrations. Once the features are identified, they can then be used to analyse differences and similarities across different types of Buddha statues. The features can be simplified to binary ‘characters’ for cladistic analysis (Marwick, 2012), or they may be modelled digitally in three dimensions so that one can appreciate their evolution by means of animations (Wisetchat, 2013a, 2013b).

Digital animation is useful in revealing the complex facial differences in what might be regarded as different ‘ethnotypes’ of Buddha statue. It is difficult to appreciate complex differences between types by inspection, where conventionally the two are presented simultaneously as adjacent illustrations and the viewer shifts gaze alternatively between the two. But if a three-dimensional model can be seen to transform continuously from the shape of one type to that of another type, where they differ they will be seen to deform. That movement draws the observer’s visual attention away from regions that are in common (and therefore unchanging). The

successful use of digital animation in the description and visualisation of Buddha statues was one of the primary motivations for this study, as many of the ideas would appear to transfer directly to considering human faces of different ethnicity.

Another motivation has to do with the challenge of creating a set of facial shape features that would be consistent with the technical conventions for describing and measuring faces, but which extends those notions into three-dimensions. This would be a novel contribution, since shape is usually left implicit — descriptions are primarily quantitative and based on measurements of dimensions between landmarks (distinguished points on faces). This motivation appeals to me personally, since the process of attempting to formalise some facial feature in terms of its shape results in my learning better how to visually appreciate the faces of people around me and better understand the clues to their ethnicity as well as individuality.

## 1.2 Research Objectives

The broad goal of this study is to create a descriptive visual representation of the human face that captures the salient characteristics common to a given ethnicity or ‘ethnotype’ (a scheme will be provided for defining these ethnotypes). A ‘descriptive visual representation’ is more than just a model or a replica. Even if a three-dimensional model were to faithfully replicate the facial features of some ethnotype, its characteristics would remain *implicit* — the model would not serve as a description any more than would a sculpture in marble or other physical medium. To be descriptive, the model must be created on a foundation of explicit features and their properties.

A set of facial attributes will be defined which constitute a vocabulary of shape descriptors. The set of attributes needs to capture sufficiently-many salient facial characteristics for each of various ethnicities. The nature of the description process (in general) is that no set would be mathematically complete. The criteria for designing a description depend on the use to be made of the description. The



attributes also need to be able to be used in arbitrary combination. Each attribute will be defined by its extremes of shape. Attribute values within that range will be represented by interpolation between those extremes.

A parametric model will be constructed capable of representing the geometric attributes in arbitrary combination. Specific ethnotypes will then be modelled by adjustment of the parameters to match the characteristic shape of that ethnotype. Ethnic differences will be visualised by interpolation between models, in essence navigating a face space between faces. Alternative navigation paths will be explored to assist in revealing ethnic differences.

The feasibility of this approach will be evaluated from several perspectives. These will include the practicality of modelling and visualising faces in terms of variations in visually-presented attributes, the utility of the resultant attribute-value pairs as a quantitative description of a specific face (or type), the repeatability and consistency with which attributes can be adjusted, and usability issues with the user interface.

It should be added that while the proposed descriptive scheme (Chapter 4) constitutes a representation of facial appearance, it is not intended to correspond to any perceptual representation employed by the human visual system when called upon to describe what we see when scrutinising a face. Consider any descriptive expertise we may possess, such as the ability to describe the flavour of wine, the design of a font face, or a melody. Each sort of description is highly cognitive, and uses a specialised vocabulary to efficiently and precisely summarise and communicate. No expectation is held that the descriptive vocabulary corresponds directly to the perceptual representations underlying olfactory, taste, vision, or audition. Likewise, a scheme for describing faces is a cognitive task. Hence when the term ‘description’ is used in this study to reveal characteristics of the face, and to distinguish differences in those characteristics, but not to correspond to any internal representation of face perception. But unlike a conventional written or spoken description using words and

phrases, the language adopted here to describe faces will be itself visual, closer to pictographic communication than text (Sections 2.3 and 9.1.4).

### **1.3 Synopsis of the Thesis**

#### **1.3.1 Literature Review**

A geographically-based definition of an ethnotype is introduced and supported by both evidence of human genetics and cultural conventions. The correlation between genetically-defined clusters and major land masses supports the basic categorisation of facial types including EUR (European), EAS (East Asian), AFR (African), and so forth. Within each ethnotype, it is commonly presumed that the population shares distinguishing characteristics that differentiate that population from other ethnotypes. The concept of an ‘average’ face is then considered in light of the considerable individual variation within any ethnic population. Anthropometry, the conventional means for measuring faces, and categorising and distinguishing among ethnotypes, is reviewed along with the traditional and more-recently-developed means to gather facial measurements. Measuring a face, however, is not the same as describing a face. The fundamental notion of creating a visual description, as opposed to a replica or depiction, is discussed, along with methods to assist in appreciation of visual differences in shape. To provide sufficient specificity, any description of the human face necessarily involves a large number of attributes. A specific face would correspond to a set of attribute-value pairs (each attribute being assigned a value as part of the overall description of that face), and this notion applies to both the description of an individual or some face representing the averaged face derived from some sample population. The space of possible faces resulting from different assigned attribute values is conventionally termed a ‘face space’. The difference between a face space intended for representing individuals within a homogeneous population and a face space for exploring ethnic differences across populations is then discussed, leading to suggestions on how to traverse that space to visualise how faces vary ethnically.

### 1.3.2 Methodology

The approach is described as one of creating a research tool as a concrete proof-of-concept demonstration of the utility of a set of ‘semantic’ (intended for human interpretation rather than automated interpretation) attributes. The modelling scheme centres on creating a set of attributes that spans the range of ethnic variation in human facial features. Given that there is no ‘average face’ on which to compare facial variation, the representation of facial attributes as interpolates between extremes of facial attribute is introduced. The general notions of implementing these ideas in terms of parametric modelling are described, followed by the specifics of the evaluation process. The working assumptions underlying this research are listed at the end of this chapter.

### 1.3.3 Facial Description

The face is naturally regarded as comprised of separate spatial regions (nose, mouth, eyes, etc.). It is fortunate that these regions permit a modular (local) approach towards describing the human face: the overall face can be described as largely the sum of the descriptions of its parts. This applies to both the description of the features within each such region, and how they combine globally to form the overall face. Individual facial regions often provide strong evidence to ethnicity. These localised features are often more subtle and fine-scaled than the point-to-point measurements taken between landmarks used in conventional anthropometry, and they often relate more to the shape of topographic features (ridges, folds, sulci, etc.) rather than their dimensions. A facial description will be regarded as a collection of *attribute-value pairs*, where the attributes are discrete, named properties of facial features, and the associated values are quantitative. Borrowing on the general tendency to base descriptions on analogies with known exemplars, the values are relative to exemplars. Some attributes will have an unsigned value (0.0 to 1.0) to represent the magnitude (or weight, or prominence, etc.) of a feature that may vary from absent (0.0) to some extreme (1.0). For other attributes, their values range between two extremes, so the value is signed (-1.0 to 1.0) and corresponds to the

fractional magnitude of this attribute within extremes (of nose width, or other attribute). All attribute values are interpolates between exemplars representing those extremes. A *model* is then comprised of a set of attribute-value pairs, which can then be used in three-dimensional visualisation, or as the basis for analysing ethnic differences.

#### ***1.3.4 Facial Features and their Attributes***

The six regions of the face (cranium, eyes, etc.) are examined individually with reference to averaged images of different ethnicities. For each facial region a set of ethnically-variable features is isolated and for each such feature its attributes or properties are identified. For example, the eye has 14 identified features, one of which, the epicanthal fold, varies from absent in some ethnicities to quite prominent in others. A total 77 facial attributes across the six regions are identified and illustrated.

#### ***1.3.5 Developing the Parametric Model***

The development of the parametric modelling environment is described, with details given of the design of the set facial features and the attributes associated with each feature. Some attributes are unsigned, varying from absent (a value of 0.0) to some extreme (1.0); other attributes are signed and vary between a minimum (-1.0) and maximum (1.0) value. The attribute value is used to interpolate between associated basis shapes which represent those extrema. The basis shapes are ‘quasi-orthogonal’, i.e., able to be added in arbitrary combination with minimal undesired interaction. The basis shapes for all 77 facial attributes are shown, region by region, implemented as three-dimensional blend shapes in Autodesk Maya®, and examples of their combinations are shown. The JSON-based file format used to represent a model is described. Finally, the matching of the resultant parametric three-dimensional model mesh with two- and three-dimensional data sources is discussed.

### 1.3.6 Ethnicity Modeller Evaluation and Applications

The *Ethnicity Modeller* (the facial attribute representation scheme and its user interface) is evaluated in terms of the *precision*, *accuracy*, and *orthogonality* of the facial attributes, and the *intuitiveness*, *efficiency*, *generality*, and *utility* of the user interface. Precision was evaluated by a repeated measurement study, showing that observed measurement uncertainty is generally less than 10% of the range of each attribute. Signed attributes that correspond to dimensions (such as mouth width or nose length) are repeatable to 5% or less, while some unsigned shape attributes (such as the depth of the shallow sulcus below the eyes or the roundness of the tip of the nose) are more difficult to judge, with a few such attributes having just noticeable differences of somewhat greater than 15%. Five dimensional attributes were calibrated in millimetres. Their mean values for two ethnotypes (East Asian and European) were compared with their corresponding anthropometric measurements and the *Ethnicity Modeller* results generally matched to within one standard deviation, i.e., the variance due to individual variation. Some problems with attribute orthogonality were identified which affect modelling accuracy and efficiency. Generality was shown in the ability to model a broad range of individuals within an ethnotype (whose features often vary more within an ethnotype than the differences of means across ethnotypes). Utility was demonstrated in the ability of the *Ethnicity Modeller* to reveal subtle and complex difference between ethnotypes through three-dimensional morphing, both interpolation and extrapolation.

## 2. Literature Review

### 2.1 Facial Types

Along with the global diversity of culture, language, ancestry, ritual, and religion, the physical characteristics of the human face vary considerably with locale. The faces of villagers in Indonesia, for example, differ from those of Scandinavian villagers. The greater the separation between two locales, the greater the facial differences.

Rather than regard human diversity as a continuum of gradual variation on many dimensions, however, there is a persistent tendency to categorise populations discretely, and to give names to these categories. Humans create national boundaries that sharply divide the global population into separate nationalities. Spoken languages similarly distinguish one population from another, often neighbouring, population. In fact the whole of human diversity is not so easily seen as a continuum as it is a consider the human world as composed of discrete set of groups or types. The groups are often (but not invariably) spatially separated, so that national, linguistic, cultural, ritual, and other boundaries, often (but not invariably) align along the same geographic borders. It is common, therefore, to distinguish (name, label) one or another population based on any one of those qualities.

The distribution of human diversity (of culture, language, physical characteristics, and so forth) is often divided into discrete types, i.e., *ethnotypes*. Now, an individual's ethnicity can often be seen in the face, despite the various ways in which individual faces differ within any one ethnicity, and even within a family. In addition to the gradual shift in shared facial features across locales (which make Indonesians appear somewhat different from Cambodians and much different from Scandinavians), there are distinctive individual differences or characteristics in the faces within any locale. This makes individual Indonesians and Scandinavians discriminable and recognisable. These ideas are of course quite familiar and intuitive. Indonesians share some common characteristics which distinguish the Indonesian face from other types. Likewise,

Scandinavians share common facial properties that dramatically distinguish them, as a group, from Indonesians and less so other closer by neighbours.

To study faces systematically, it is traditional to impose a classification scheme on the overall population of human faces. But for any such set of discrete types (e.g., African versus Asian) there may be blends of these types, such as Afro-Asian. No single set of facial categories has yet been broadly adopted, but many ad hoc categories are readily derived by choosing larger or smaller geographical distinctions (such as distinguishing South, versus Southeast, versus East Asian rather than just Asian). The choice of categorisation scheme depends on the uses to which it will be applied. Ultimately, human ethnic variation, just like genetic variation, is effectively a continuum created by the billions of humans now alive and the sometimes intertwined, sometimes separated histories of their ancestors. Nonetheless, there is value to using some categorisation scheme to create (artificial but convenient) subdivisions in this continuum.

### *2.1.1 Categorising Ethnotypes Geographically*

There is long history of attributing ethnicity based on geography. Each ethnotype has a distinct geographical centre, and often named after a region such as Caucasoid or Mongoloid. Ethnicities have long been distinguished on the basis of differing physical and cultural characteristics. Early categorisations by Giordano Bruno (1548-1600) and Jean Bodin (1530-1596) were based on skin colour, then other characteristics such as stature, shape, and cultural differences were used (e.g., by Bernhard Varen (1622-1650), John Ray (1627-1705), and François Bernier (1625-1688)) to map different ethnicities around the globe. Bernier distinguished four groups (Europeans, Far Easterners, Negroes, and Lapps) distributed around the “four quarters of the globe” (Bernier, 1684). A century later, Johann Friedrich Blumenbach (1752-1840) proposed five races based on cranial measurements (Blumenbach, 1795). The English biologist Thomas Huxley (1825-1895) distinguished nine racial categories on the basis of appearance and anatomy and mapped them geographically (Huxley, 1870). While Huxley’s terminology is no longer fashionable, the basic classification scheme he laid out is largely still in use. A

subsequent map of races published in the Meyers Konversations-Lexicon (1885-1892) distinguished the *Caucasoid* race (consisting of Aryans, Semitics, and Hamitics), while the rest of the world's ethnicities were basically classified as *Negroid* or *Mongoloid*. A growing European-centric elitism regarded the Caucasians as the most advanced of the three races; Huxley (1870) regarded it an “absurd denomination” to group all Europeans into a single Caucasian race. While many 19<sup>th</sup> century naturalists proposed categorisations that split the human population into separate or ‘pure’ races or even species, Darwin (1859) noted “.. that [races] graduate into each other, and that it is hardly possible to discover clear distinctive characters between them”. Eventually, the defining of pure races gradually fell out of favour. Following World War II, the United Nations concluded that it would be better when speaking of human races to drop the term ‘race’ altogether and speak of ethnic groups (Metraux, 1950).

While the term ‘race’ has been replaced by ‘ethnic group’ (or ‘ethnicity’), the practice of using geography to distinguish among these groups persists (e.g., ‘Asians’ versus ‘Africans’ in broad terms, and ‘Nordic’ versus ‘Mediterranean’ types on a smaller geographic scale. Sufficiently large ethnic groups (the Han Chinese, for example) are readily divided into subgroups (the Cantonese, for example, is a subgroup of the Han Chinese ethnic group). The concept can therefore be broadened or narrowed, from regarding all Asians as one group to distinguishing East Asians, to only Chinese, to Han Chinese, and so forth. Recent studies of the human genome are validating this simple and intuitive approach towards distinguishing ethnotypes.

### 2.1.2 *Categorising Ethnotypes Genetically*

The 19<sup>th</sup> century view that races correspond to discrete types separated by distinct gaps — the ‘typological’ view of race as termed by Ernst Mayr (1962) — is not supported by recent studies of human genetics. That is, no ‘race gene’ has yet been found that is exclusively present in one race and absent in others. An alternative to the typological view, the ‘populationist view’, suggests that races differ from one another statistically, not absolutely or categorically (Mayr, 1962). This view is supported by the study of the



human genome: there is a continuous gradation in gene frequencies, where each population blends into another (Cavalli-Sforza and Menozzi, 1994; Sarich and Miele, 2003). Some genes are more prevalent in some populations than in others; the human population is actually a *continuum* of ‘types’. Sarich and Miele (2003) conclude that races are “... supposed to blend into one another, and categories need not be discrete. It is not for us to impose our cognitive difficulties upon Nature; rather we need to adjust them to Nature”. Modern genetic studies support the populationist view, and refute the typological view of races as discrete and separable types. So while there are no ‘pure’ races, there might still be a genetic basis for distinguishing races based on the statistical distribution of genes across populations.

Nearby populations tend to resemble each other more than distant ones, but within each population there is very large variability. Two individuals from different populations could often be more genetically similar than two individuals from the same population (Bamshad et al., 2004; Witherspoon et al., 2007). Nonetheless, genetic differences accurately assign most individuals to the correct population of origin despite the high diversity within each population (Witherspoon et al., 2007).

While the boundaries between races (whatever they are) are not definite, “... there appear to be six main genetic clusters, five of which correspond to major geographic regions, and subclusters that often correspond to individual populations” (Rosenberg et al., 2002). The similarity and dissimilarity in DNA sequences sampled from individuals, subjected to cluster analysis, reveals a significant geographic influence. If the world’s population were divided into only two groups, they correspond to the two major land masses: 1) Europe, West Asia, and Africa, and 2) East Asia, Australia, the Americas. Significantly, if five genetic groups are distinguished, they comprise: 1) Europe plus West Asia, 2) Africa, 3) East Asia, 4) Australasia, and 5) the Americas. It bears remembering that two centuries earlier, Johann Friedrich Blumenbach (1795) proposed essentially the same five categories: Caucasian, Negroid, Mongoloid, Malay, and Native American.

Rather than using the term ‘race’ or ‘ethnic group’, geneticists frequently adopt the term ‘population’ as a distinct group of interbreeding people, often associated with geography (Lewontin, 1972). The earlier idea of discrete racial types (each with a sharp boundary) is therefore being replaced with the idea of populations or clusters. These populations are each quite diverse internally and they overlap, especially when populations are geographical neighbours. Nonetheless, the centres of these clusters correspond to geographically-separate regions. While the focus has shifted to understanding these clusters genetically, and human populations are now being categorised in terms of haplogroups, there is significant geographic order in the distribution of haplogroups.

The early notion of ‘race’ has gradually been replaced with the notion ‘ethnic group’. Ethnic groups (or ‘populations’) are often associated with the region of origin for those of a given ethnicity. Geographic region, of course, is only one means by which ethnic groups are distinguished, as they also have distinct shared cultures, religions, customs, and so forth. Maps of genetic diversity correlate well with traditional maps of ethnicities (e.g., by Huxley, 1870), but the sharp boundaries drawn in such maps, while artistic and tidy, do not reflect the very gradual and overlapping nature of actual genetic distributions.

While hard boundaries should not be expected between ethnicities, the continents themselves form natural boundaries for many ethnicities. Other ethnic groups are surrounded by neighbouring ethnicities, and while the geometric extent of the given group (such as Mongolic) may be indistinct, there is often a sense of a centre (corresponding to the highest concentration of that given ethnicity). In some cases national boundaries are ethnic boundaries. Island nations such as Japan or nations on peninsulas (e.g., Korea), and the nations of Western and Northern Europe, are among some of the most ethnically homogeneous, while those of Africa and the Middle East are the most ethnically diverse (Alesina et al., 2002). Across a larger geographic scale, however, the African populations, despite their broad ethnic diversity within each nation, share more ethnic similarity than those of countries on other continents. The continental

boundaries, therefore, effectively segregated the human populations geographically and ethnically through human history, as revealed by human genetics by the five main clusters: 1) Europe and West Asia, 2) Africa, 3) East Asia, 4) Australasia, and 5) the Americas.

United Nations Statistics Division (2007) devised a geography-based for ethnic regions and subregions and member states. Five basic geographical regions are defined: Africa, the Americas, Asia, Europe, and Oceania, each with subregions and each subregion with member states. A purely geographic definition of ethnicity works well to distinguish broadly between the African versus Asian versus European ethnotypes. Geography is not a straightforward indicator of ethnicity, however, in regions where the indigenous or aboriginal populations are a minority. The native populations of the Americas (North, Central and South America) were greatly overshadowed by the post-Columbian influx of primarily Europeans and Africans. Similarly, the ethnicity in Oceania has been greatly complicated by colonisation.

### *2.1.3 The Risk of Stereotyping*

While only valid in broad strokes, it is convenient to adopt a straightforward geographic categorisation of ethnotypes for the current study, following the UN classification scheme (EUR, SEA, etc.). Each region is further broken down to subregions to which a three-letter code is added (e.g., East Asian is EAS). While the purely-geographic UN approach reflects many ethnic hierarchies rather well (e.g., JP is in EAS, which is within ASN), it was noted that some ethnicities are indigenous minorities within a larger current population that reflects a large colonial influx.

The convention of assigning ethnotypes geographically is intended only for the scope of this study. Due to the overlapping nature of ethnic variation, it would be simplistic to attempt to can draw a discrete boundary around a region such as SEA or EAS. One cannot delineate an ethnotype so neatly. In the rest of this thesis, therefore, a labeled ethnotype (such as EUR) is intended only to compare different population samples, in the

same manner that Farkas et al. (2005) compare anthropometric measurements across separate samples.

In applying any categorisation scheme, there is a risk of stereotyping. Statistical methods in general based on measurements of a sample of some population raise this potential when the mean of some property is falsely presumed to apply to all individuals. Later in this chapter two very different statistical methods will be reviewed (Section 2.2). One is to create an average image composed of the images of multiple individuals, intended to summarise the appearance of an average face for some ethnicity. The other statistical method, anthropometry, creates averages of dimensional measurements taken of multiple individuals. In studying the face characteristics of some ethnicity, say EUR, sample statistics may be used in an overly simplistic manner, wherein the dimensions, proportions or appearance of a sample of EUR individuals is misconstrued to be stereotypic of all EUR faces, however. As will be shown, face statistics (from either image averaging or measurements) carry with them evidence of considerable variance in most metrics, and skew, rather than normal, distributions. Caution is therefore needed to avoid over-generalising about the facial characteristics of different ethnicities.

Another caution is in order. While this study introduces a new scheme for explicitly describing facial variation, that scheme has not yet been used to actually describe ethnic variation (only a pilot study is demonstrated in Section 7.2.1 with details in Appendix 3.4). Nor has the scheme been adopted to a very different purpose such as automated ethnicity recognition (an extension of face recognition).

#### ***2.1.4 The Concept of an Average Face***

Conceptually, the ‘average’ face represents the common or dominant characteristics of a population and hopefully disregards the individual facial variations within that population (and the contribution of differing subpopulations). The face of an average East Asian, for example, would reflect contributions from the various EAS populations (primarily Chinese, Japanese, Korean, and Mongolian). But how should the different

contributions to EAS be weighed? The total population of EAS is roughly 1,600,000,000. Weighted by relative population size, the average EAS face would be essentially a Han Chinese face since about 88% (about 1,400,000,000) of EAS are Chinese, and of that vast majority, about 90% are more specifically Han Chinese. While the Chinese ethnotype greatly outnumbers the remaining roughly 180,000,000 EAS populations (composed of primarily 127,000,000 Japanese, 51,000,000 Korean, and about 3,000,000 Mongolians), those other ethnotypes have their own distinctive combinations of facial features compared to the ‘average’ EAS (United Nations Department of Economic and Social Affairs and Population Division, 2015).

Fortunately, studies of ethnic facial variation are not bound to population-weighted averages, of course. Of greater importance in the study of facial variation is to understand the distinctive characteristics of the face that vary with ethnic population regardless of the population size. A small yet homogeneous sub-population such as Mongolian or Japanese may present salient differences from the dominant population of Han Chinese facial types.

In comparing populations statistically, the presumption is that facial measurements are distributed normally, and with sufficiently-small variation around their respective measured means. The concept of an average face is thus decreasingly useful as the population is increasingly diverse. To find the mathematical mean for some property the (often tacit) assumption is that the underlying population is normally distributed. But by including individuals from multiple distinct subpopulations, the resultant sample distribution might be skewed from normal. Even if the distribution is not obviously multi-modal, the large aggregate variance might render the mean an imprecise and not very useful measurement. In particular, it is problematic to classify as AFR the entire population of the African continent. An aggregate population of about 1,200,000,000 AFR people distributed over about 30,000,000 square kilometres of the African continent (about 20% of the Earth’s land area) should not be expected to comprise one homogeneous, normally-distributed population. Nonetheless, AFR is sometimes

regarded as a single ethnotype, as in studies that compare African American males with North American Caucasian males (Porter, 2004; Farkas et al., 2005). Despite the very heterogeneous origins of African Americans (and North American Caucasians, for that matter), the considerable differences between these two populations does result in many comparisons reaching statistical significance in those studies. Recognising the considerable diversity of AFR faces, sub-classification schemes have been introduced (Ofodile et al., 1993; Porter and Olson, 2001). Any ethnotype can be subdivided into subcategories, such as EUR being regarded as more specifically composed of northern European (NEU), western European (WEU), and so forth.

The idea of an ‘average face’ for large populations such as EAS and AFR is therefore suspect. But even a population as ethnically homogeneous as the Japanese, distributed over a relatively small geographical area, may show large variation (Section 2.2.1). The complex nature of the population statistics within various ethnotypes has not discouraged researchers from laboriously measuring and comparing faces from various ethnotypes.

## 2.2 Measuring Faces

An individual’s ethnicity (i.e., the broad geographical region of an individual’s origin, if not subregion) can often be distinguished at a glance. Ethnicity is revealed by the colour of the skin, hair, and eyes, and the texture of the hair and skin and by facial dimensions, shapes, and proportions characteristic of each ethnicity. The expectation is that nose width, eye separation, and so forth vary measurably across ethnicities, on average. The specific measurements of an individual’s face would be expected to fall within some distribution typical of their ethnicity, however facial shape and proportion varies significantly among individuals of a given ethnicity, and not just across ethnicities. Hence the specific measurements taken of two African faces will not be identical, nor would those of two Asians, but the *average* measurements for Africans and Asians would nonetheless be expected to be statistically distinguishable, for some appropriate set of measurements. Two applications of anthropometry are therefore suggested:

distinguishing individuals within a given ethnicity, and distinguishing ethnicities based on averaged measurements.

### 2.2.1 Facial Landmarks and Measurements

Facial dimensions such as the width of the nose are measured between anatomically-defined locations termed *landmarks*. The distance between any pair of landmarks may be measured, providing a factor that might be correlated with ethnicity, or just reflect individual differences. The specific choice of anthropometric measurements becomes important, since a set which distinguishes the identity of an individual (from others of the same ethnic group) may not be well suited for distinguishing the ethnicity of that individual.

Anthropometric measurements of the face are based on a set of specific anatomical locations, or *landmarks* (Brothwell and Harvey, 1965; Farkas, 1981, 1994). A few common landmarks are given in Table 2.1 and illustrated in Figure 2.1. Landmarks are measurable superficially on the soft tissue (e.g., Farkas et al., 2000) and many correspond to underlying osteological features (Swennen et al., 2006).

Most landmarks are defined by the extremes of facial skin structures, such as folds, commissures, sulci, and prominences. Nose width, for example, is measured between the extremes of the left to right nostrils or alare (*al*), which is often simplified to *al-al*, and eye separation can be measured by the distance between the inner corner (endocanthion *en*) of the left and right eyes. Such absolute measurements (typically measured in mm) permit quantitative comparisons between individual faces, or of a given face relative to some ideal, or to accumulate statistics within a population to quantify differences across populations. While facial measurements are traditionally made with callipers or other physical instruments, digital imaging has been shown to provide comparable accuracy for anthropometry (Stephan et al., 2005; Sforza and Ferrario, 2006) despite some limitations (Allanson, 1997; Douglas et al., 2003).

Table 2.1. The anatomically-defined facial landmarks shown in Figure 2.1.

Landmark	Definition
<b>Midline landmarks</b>	
<i>Gnathion (gn)</i>	Point on midline of most anterior projection of the chin
<i>Nasion (n)</i>	Midpoint on the bridge of the nose; deepest point below the glabella
<i>Subnasion (sn)</i>	Midpoint of the base of the nose where the columella meets the philtrum
<b>Paired landmarks</b>	
<i>Alare (al)</i>	The most lateral point on each alar contour (lateral edge of nostril)
<i>Cheilion (ch)</i>	The point located at each labial commissure (corner of the mouth)
<i>Endocanthion (en)</i>	Inner commissure of each eye fissure (inner corner of the eye)
<i>Exocanthion (ex)</i>	Outer commissure of the eye fissure (outer corner of the eye)
<i>Gonion (go)</i>	Most lateral point on the mandibular angle (corner of the jaw)
<i>Tragion (t)</i>	Point on the upper margin of each tragus (of the ear)
<i>Zygion (zy)</i>	Most lateral point of each zygomatic arch (most lateral point on cheek)

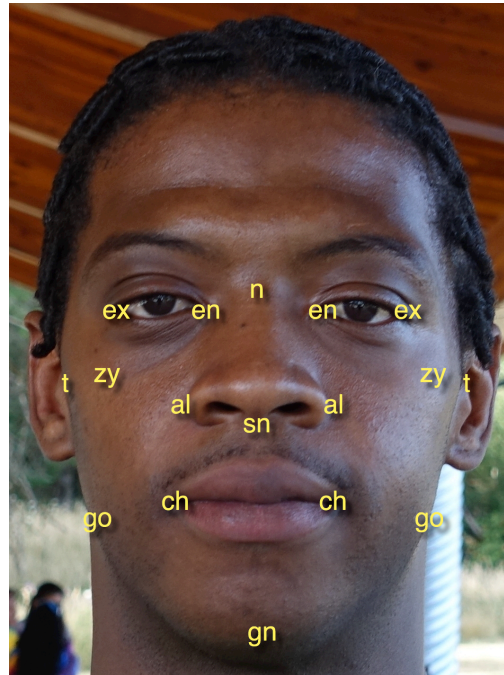


Figure 2.1. The locations of the anthropometric landmarks defined in Table 2.1.

Anthropometric studies usually measure only certain point-to-point distances and angles (Farkas et al., 2005). Traditional anthropometric research appears to have been influenced by both what can be reliably measured and what measurements seem intuitive. Since only certain anatomical landmarks can be reliably located on the surface



of the skin (Swennen, 2006), the conventional choices for anthropometric factors often reduce to the intuitive measurements, such as nose height and width.

*Absolute Measurements:* Anthropometric measurements are also commonly compared between faces of different ethnicities in order to quantify how the dimensions of the nose, or the slope of the forehead varies with ethnicity (Brothwell and Harvey, 1965; Porter and Olson, 2001; Choe et al., 2006; Tuncel et al., 2013).

In a large study, Farkas et al. (2005) provided a detailed compilation of 14 superficial (rather than osteological) anthropometric measurements across a broad range of ethnicities (26 countries over five regions of the world), for both sexes, all compared with EUR (or more specifically, the white American) face.

As expected, the overall size of the human face varies significantly across ethnicities as well as between sexes. African males, on average, have greater face height than European males (*e.g.*, *n-gn* is about 12% greater) (Farkas et al., 2005). A tabulation of some absolute size differences is shown in Table 2.2.

*Table 2.2. A selection of absolute anthropometric measurements, in mm, derived from (Farkas et al., 2005). Males from two broad ethnicities (European and African) and three Asian ethnicities (Indian, Thai, and Japanese) are compared. Each mean and standard deviation is based on 30 samples.*

Measurement	EUR_M	AFR_M	IN	TH	JP
<i>al-al</i>	34.7 ± 2.6	44.1 ± 3.4	37.9 ± 3.5	40.8 ± 2.2	38.2 ± 2.5
<i>ch-ch</i>	53.3 ± 3.3	54.6 ± 4.1	51.0 ± 4.6	50.3 ± 2.4	48.4 ± 3.5
<i>en-en</i>	32.9 ± 2.7	35.8 ± 2.8	34.1 ± 2.2	37.2 ± 3.0	37.5 ± 3.1
<i>en-ex</i>	31.2 ± 1.3	32.9 ± 1.6	30.2 ± 2.0	29.8 ± 2.0	30.7 ± 2.3
<i>ex-ex</i>	89.4 ± 3.6	96.8 ± 4.6	98.8 ± 3.5	91.5 ± 4.5	103.9 ± 4.2
<i>n-gn</i>	121.3 ± 6.8	125.9 ± 8.2	112.5 ± 5.7	123.5 ± 6.0	122.8 ± 6.9
<i>n-sn</i>	53.0 ± 3.5	51.9 ± 3.0	47.2 ± 3.7	51.5 ± 2.2	56.9 ± 4.9
<i>zy-zy</i>	137.1 ± 4.3	138.7 ± 5.6	135.8 ± 4.3	147.1 ± 5.5	147.2 ± 5.6

A few commonly-observed generalisations can be drawn. Africans have shorter and wider noses, wider faces, and taller foreheads than Europeans, for example (see also Porter and Olson, 2001). Generally, there is greater inter-ethnic variability in horizontal (lateral) measurements compared to vertical absolute measurements.

Most absolute measurements are not different statistically across ethnicities (Farkas et al., 2005). For instance, the mean face width  $zy-zy$  for a sample of 30 Africans American males is  $138.7 \pm 5.6$  mm compared to  $137.1 \pm 4.3$  mm for white North American males. This 1.6 mm difference in means is not statistically significant (Farkas et al., 2005, table 26a). In fact, face width  $zy-zy$  is one of the least variable dimensions of the human face across ethnotypes (Jeffries et al., 1995; Fang et al., 2011), while  $al-al$  and  $en-en$  among the most variable.

In broad terms, few of the characteristic features that distinguish different ethnotypes, while visually salient, are reliably revealed quantitatively by absolute measurements. The concept of an average face is further weakened by the observation that even if a population is ethnically quite homogeneous, as are the Japanese, individual variations among Japanese are very large numerically (see Tables 2.2 and 2.3), in fact comparable to the statistical variance measured in samples from an aggregation as heterogeneous as all of AFR.

While *absolute* measurements between landmarks are therefore of limited use for distinguishing ethnicities, they will have important application towards model calibration. Next, the extent to which the *relative* proportions within the face reveal ethnic differences is discussed.

### 2.2.2 Anthropometric Indices and Facial Proportions

Any set of absolute measurements by itself does not make explicit the common traits or properties that are associated with the faces in any locale, nor the differences among their individuals. While the distances between landmarks vary from individual to individual, especially with the overall size of an individual face, mathematical quotients or ratios of

such measurements are usually computed to create relative measurements. These provide proportions (such as nose width relative to nose height) which are less dependent upon the absolute size of the individual. These proportions or ratios are termed *indices*. The practice, fundamental to anthropometry, is to compare *quotients* of absolute measurements. Quantifying facial proportions in terms of ratios of absolute measurements has a long tradition in anthropometry, as discussed next.

*The Neoclassical Canons:* One of the earliest uses of facial measurements was to judge facial proportions. A quotient of two absolute measurements lead to a ratio or proportion. Early studies of ideal facial proportions by da Vinci and Dürer proposed idealisations, or ‘neoclassical canons’, such as ‘*the nose width equals the distance between the eyes*’ and ‘*the distance between the eyes equals the width of each eye*’ (Dürer, 1981; O’Neill and da Vinci, 1983).

The measurements for such canons are made between anatomical landmarks, such as the inner and outer corners of the eye (*en* and *ex*), as discussed with a representative listing provided in Table 2.2. A modern set of proportionality rules, or expectations, has been compiled as “The North American Caucasian standard” (Farkas, 1994) and expressed as a set of equalities and inequalities (Farkas, 1994). Those canons for which anthropometric measurements between landmarks are listed in Table 2.2 include:

The Orbital Canon:  $en-en = ex-en$

The Orbito-nasal Canon:  $en-en = al-al$

The Naso-facial Canon:  $al-al = 0.25 (zy-zy)$

The Naso-oral Canon:  $ch-ch = 1.5 (al-al)$

These are all ‘horizontal’ rules; others are proposed to describe vertical proportions. While the neoclassical canons form expectations for the ‘ideal’ facial proportions (e.g., Chandra et al., 2012), they most closely correspond to the proportions of European faces, compared to those of African or Asian faces (e.g., Farkas et al., 1985; Dawei et al., 1997;

Sim et al., 2000; Le et al., 2002; Porter, 2004; Fang et al., 2011; Jayaratne et al., 2013) and revisions suggested for non-EUR ethnotypes (e.g., Farkas, Forrest, and Litsas, 2000).

In the search for a basis for describing and comparing faces, the neoclassical canons can either be regarded as a failure of sorts (when they do not apply as universally as originally suggested), or constructively as a basis for comparison across ethnotypes. While comparative studies across ethnotypes often report absolute differences of anthropometric measurements, a neoclassical canon that is originally expressed as an equality can be analysed in terms of inequalities. For example, Porter and Olson (2001) showed that the Orbito-nasal Canon ( $en-en = al-al$ ) does not apply well to AFR females ( $en-en < al-al$  for 93.5% in a sample set of 108 individuals). The neoclassical canons in general are now recognised to have limited validity as general standards which apply (at best) only for EUR, but they demonstrate two useful ideas of facial description: 1) attempting to find size-invariant descriptors, where a proportion created by a quotient normalises one absolute measurement relative to another, and 2) describing by comparison to a standard, i.e., by analogy.

*Categorical Distinctions:* A few indices have long been known to distinguish broad ethnic categories, such as the variation in *nasal index* (the ratio of nose width to height, multiplied by 100, by convention). Based on skull measurements, the 1858 edition of Gray's Anatomy (Gray, 1918) defined the narrow *leptorrhine* noses of Europeans (index  $< 48$ ), the *mesorrhine* noses of Asians (48-53), and the broad and flat *platyrrhine* noses of Africans and Australian Aborigines ( $> 53$ ). This same practice was applied to other indices, such as the cephalic, orbital, and gnathic indices, to classify other skull proportions (Duckworth, 1904). But this practice of categorising faces based on index ranges has limited utility beyond the very broad distinctions recognised over a century ago (Franco et al., 2013). While categories may not be very useful for distinguishing ethnicities, indices do quantify facial proportions, which do vary across ethnicities.

*Indices and Ethnicity:* Computing indices is intended to remove the effect of overall facial size (since one is interested in the proportion of nose width to nose height, the nasal index, for example). The seven anthropometric indices in Table 2.3 are derived from quotients of the corresponding absolute measurements in Table 2.2. Each index is the quotient of two statistical absolute measurements, and therefore subject to error propagation (Taylor, 1997). Those indices are computed for broad ethnicities EUR, AFR, and to similar ethnicities JP and TH, then compared pairwise between ethnotypes (EUR vs. AFR, AFR vs. JP, and JP vs. TH).

In the case of the nasal index ( $al-al/n-sn$ ), African and European noses are significantly different, as are Japanese versus Thai noses. The majority of other indices are not significantly different, however; see (Farkas et al., 2005) for a more detailed discussion of how few anthropometric measurements are significantly different across these ethnotypes.

If the goal were to distinguish simply AFR versus EUR or AFR versus EAS, various indices involving nose width  $al-al$  might well suffice, such as  $al-al/en-en$ ,  $al-al/n-sn$ , or

*Table 2.3. Anthropometric indices computed from (Farkas et al., 2005), with standard deviations for each quotient computed by error propagation of the sample means (Taylor, 1997). Males are compared for European versus African, African versus Japan and Japan versus Thai. Some comparisons are provided across ethnicities by unpaired  $t$  test. NS indicates not significantly different.*

Index	EUR	AFR	JP	TH	EUR v. AFR	AFR v. JP	JP v. TH
$al-al / en-en$	$1.05 \pm 0.12$	$1.23 \pm 0.14$	$1.02 \pm 0.11$	$1.10 \pm 0.11$	$p < 0.0001$	$p < 0.0001$	$p = 0.0066$
$al-al / n-sn$	$0.65 \pm 0.07$	$0.85 \pm 0.08$	$0.67 \pm 0.07$	$0.79 \pm 0.05$	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$
$ch-ch / zy-zy$	$0.39 \pm 0.03$	$0.39 \pm 0.03$	$0.33 \pm 0.03$	$0.34 \pm 0.02$	NS	$p < 0.0001$	NS
$en-en / ex-ex$	$0.37 \pm 0.03$	$0.37 \pm 0.03$	$0.36 \pm 0.03$	$0.41 \pm 0.04$	NS	NS	$p < 0.0001$
$en-en / zy-zy$	$0.24 \pm 0.02$	$0.26 \pm 0.02$	$0.25 \pm 0.02$	$0.25 \pm 0.02$	NS	$p = 0.058$ NS	NS
$zy-zy / n-gn$	$1.13 \pm 0.07$	$1.10 \pm 0.08$	$1.20 \pm 0.08$	$1.19 \pm 0.07$	NS	$p < 0.0001$	$p = 0.61$ NS
$al-al / zy-zy$	$0.25 \pm 0.02$	$0.32 \pm 0.03$	$0.26 \pm 0.02$	$0.28 \pm 0.02$	$p < 0.0001$	$p < 0.0001$	$p = 0.0031$

*al-al/zy-zy* (Table 2.3), from (Farkas et al., 2005). However the large individual variation observed *within* each ethnotype, and the small differences in means *across* ethnotypes casts doubt on this approach for all but the broadest differences across ethnicity (e.g., AFR versus EUR). Larger sample sizes will not solve that problem.

An alternative to calliper-based anthropometry, the study of faces through image analysis, discussed next, not only provides more detailed quantitative distinctions across ethnicities, but also a path towards understanding how best to visualise those similarities and differences.

### 2.2.3 *Measurement Methods*

Recognising that facial differences are difficult to quantified by a few statistical measurements taken by callipers, techniques have been developed to capture the face as a whole. The first approaches were purely photographic (film-based) and used image superposition to create an average across multiple individual faces (Galton, 1878). Subsequent refinement of those methods using digital image processing has permitted quantitative reconstructions of both individuals and averages of sets of individuals. Regarded as sets of measurements, these techniques result in far larger datasets than can be practically achieved using manual measurements. Moreover, they permit visualising the ‘average face’ of some sample set, and not merely a tabulation of statistics.

*Image-Based Measurements:* The concept that a given population has an ‘average face’ underlies conventional anthropometry. The population is presumed to be homogeneous and that each set of facial measurements is normally distributed about its respective mean. Those measurements then are used to quantify the ‘average face’, usually for a small population sample, and for a relatively small number of such anthropometric measurements. Some influential studies involved only 30 or so individuals (e.g., Farkas et al., 2005).

An ‘average face’ can also be created by superimposing the images of multiple faces. While the population sample may remain small, image averaging provides a more holistic representation of the face. The concept of averaging images originated with the ‘composite portraiture’ of Sir Francis Galton (1878). Galton was unsuccessful in his attempt to detect stereotypes (e.g., the face of the chronically sick or the criminal face) by compositing examples of such individuals, but was successful in discovering that the averaged face depicted by a composite of multiple superimposed faces was more attractive than that of any of the individuals making up the composite (Benson and Perrett, 1991). As discussed in Section 2.1.3, image averaging can readily lead to stereotyping, especially when the underlying statistical variance in the sample set is ignored. As image registration techniques have improved with digital image averaging, facial features are increasingly well preserved (e.g., DeBruine and Tiddeman, 2017). Nonetheless, averaging tends to create a misleadingly simple picture of the ‘average face’, as discussed momentarily, which could lead to overly-simplistic conclusions.

Composited facial images provide two advantages over physical callipers-based anthropometry: they allow one to visualise the shape of the average face as a whole, not just numerical values between discrete landmarks (such as *al-al*), and they also provide far more measurements than would be practical with physical techniques.

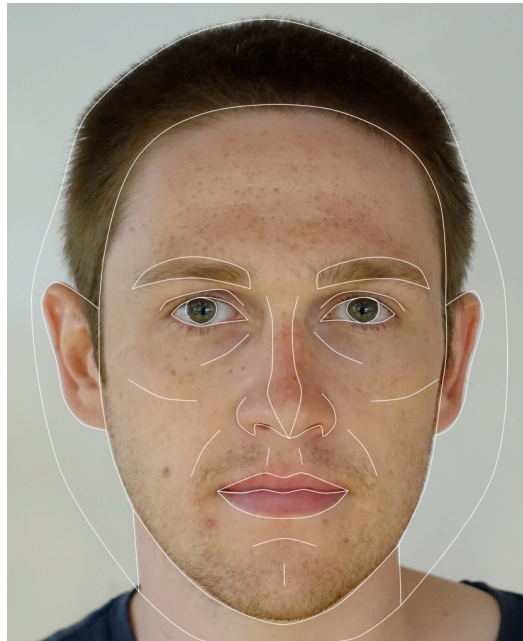
The shape of a face (and not just its numerical dimensions and overall proportions) is well-recognised to be important for facial recognition (Stephan et al., 2005). While facial measurements and proportions of course quantify some aspects of the shape of a face (narrow nose, wide mouth, etc.), those measurements do not directly or explicitly address the specific shape of the nose or mouth. For the shape of a facial region to be measured numerically would require far more than the small number of point-to-point measurements typically gathered by callipers and delineation of two-dimensional images. Fortunately, image averaging provides both shape information for us to appreciate visually as well as measure. But it must be stressed that the shape of a face remains implicit in an image. It is still just a depiction.

The creation of a composite image requires spatial registration of the individual images, through a process of delineation, followed by a process of image superposition to create an average image (Langlois and Roggman, 1990; Benson and Perrett, 1993; Perrett et al., 1994; Tiddeman et al., 2001; Stephan et al., 2005). Importantly, image averaging not only creates an averaged image, but also provides for a greater number of anthropometric measurements, due to the larger set of landmarks measured on each face, compared to the laborious use of callipers. Both techniques are subject to measurement error (Pérez-Pérez et al., 1990), but image averaging and physical measurements have been shown to have similar accuracy when affects such as image scale and perspective are carefully controlled for (Stephan et al., 2005; Ghoddousi et al., 2007).

*Delineation:* Delineation is the process of mapping the location of specific facial landmarks or ‘fiducial points’ such as the location of the endocanthion in the image. To perform delineation, a template (essentially an outline drawing with straight lines connecting the fiducial points) is superimposed on the image, and then the position of each fiducial point is aligned with the corresponding facial feature (Figure 2.2). For example, delineation in Webmorph (DeBruine and Tiddeman, 2017) begins with mapping three fiducial points: the left and right pupils plus the stomion. Based on these three points, the scale, orientation, and placement of the overall template is adjusted automatically, and one makes fine adjustments to the individual fiducial point. Once delineated, the resulting template file can be thought of as capturing the shape of a given face as a file of image coordinates. The delineation process is then repeated for each image in the set of images to be averaged.

*Averaging:* Given a set of template files created by delineating a set of face images, the next step is to create an average template file by computing the mean position for each fiducial point within each template, across the set of templates to be averaged. The resulting average template can be regarded as the delineation of the average face. Since some of these fiducial points correspond to conventional landmarks, this process





*Figure 2.2. Image with superimposed delineation. The white outlines trace a total of 189 defined fiducial points, some of which correspond to classical landmarks, others used to assist in bounding regions such as the eyebrows for subsequent image averaging.*

provides an alternative source of information for statistical means of anthropometric measurements such as *al-al*, and *n-sn*. Given that a template may consist of about 200 fiducial points, image-based anthropometry allows more detailed facial measurements than would be practical using physical callipers.

With the averaged template providing the mean location of each fiducial point, the next step is to fill in the averaged colour and texture across the face by defining a triangular mesh formed by neighbouring triples of fiducial points, and filling in those triangular patches. Since all templates have the same mesh topology, each triangle across the averaged template has a corresponding triangle in every individual template. The colour and texture of each individual triangle can therefore be combined to create the averaged colour and texture in the resulting average image (Benson and Perrett, 1993; Perrett et al., 1994; Tiddeman et al., 2001). The end result of image averaging is then a delineated image, where each fiducial point is the mean of the corresponding fiducial points from



*Figure 2.3. Photographs of 11 AFR females were delineated and averaged to create the image AFR\_F11 (centre image, upper left). The mean template (lower left) does not reflect the large individual differences that are apparent when the individual templates are superimposed with a common interpupillary distance. Photographs by L. DeBruine, Face Research Laboratory, University of Glasgow, and the author.*

the sample images, and the shading and texture across the averaged face is the average of that across the sample images (see Figure 2.3).

An averaged image will have sharply delineated features such as the outline of the eyes, and yet the texture and shading will appear to vary smoothly in the regions between the delineations. As shown in the individual templates of Figure 2.3, however, the averaged

image does not reveal the sometimes-considerable variation among the individuals that were averaged. The superimposed delineations in the lower right are a graphical reminder of the large statistical deviations (on the order of many millimetres) of within-ethnotype measurements reported in anthropometric studies (Farkas et al., 2005 and Table 2.2).

In Figure 2.3 the centres of the left and right pupils are all superimposed, which essentially normalises all faces to have constant IPD. Since IPD is, of course, not constant across individuals or ethnotypes (Moorees et al., 1976), it would be preferable to have image scale held precisely constant across all photographs and to use one reference point such as the midline point of the glabella (Stephan et al., 2005). This would avoid the illusion that face width and other dimensions vary when in fact IPD was varying.

*Transforming:* A delineated image can be morphed or warped by moving the fiducial points of its template. This can be a powerful tool for manipulating facial images. For example, given two delineated face images  $I_S$  and  $I_D$  (for source and destination), a transformed image  $I_T$  can be created that is a fractional blend (or interpolation) between  $I_S$  and  $I_D$ . The transformation process involves first computing a new template to be associated with the transformed image  $I_T$ , in which the position of each fiducial point is interpolated to be a fraction of the way between the location of that point in  $I_S$  and in  $I_D$ . The interpolated template specifies the geometry of  $I_T$ . The colour and texture are then interpolated blends between those of the images  $I_S$  and  $I_D$ . Transformation can be used for ‘morphing’ to simply interpolate from one image to another, or to compute a difference between two images and apply that difference to a third image (Tiddeman et al., 2005). For an interpolation parameter  $\alpha$  (where  $0.0 \leq \alpha \leq 1.0$ ), the transform would be called *morphing* if  $I_T = I_S + \alpha(I_D - I_S)$ . One can also create exaggerations by *extrapolating* rather than interpolating, i.e. using values of  $\alpha > 1.0$  in order to create a caricature (Benson and Perrett, 1991; McIntyre et al., 2013). Also, one can apply the difference between  $I_D$  and  $I_S$  to a third image  $I_O$ :  $I_T = I_O + \alpha(I_D - I_S)$ . For example, if  $I_D$  and  $I_S$  are images of the average male and average female, their difference ( $I_D - I_S$ )

captures the geometrical differences between the two sexes. One can, for example, then make the image of an individual androgynous, or an female appear more masculine, or to exaggerate the masculinity of some male face (Rennels et al., 2008; Scott et al., 2010; Holzleitner et al., 2014). This technique has been applied to studies of face recognition, ageing, symmetry, attractiveness, beauty, and cosmetic surgery (Burt and Perrett, 1995; Perrett et al., 1999; Chandra et al., 2012; Lam et al., 2013; Coetzee et al., 2014).

Since the transform data is provided *only* by the set of discrete fiducial points, there is no information about shape change in the regions *between* these sample points. Texture and colour must be interpolated to appear continuous in the regions away from the fiducial points. While interpolation of colour and texture pixels may result in a smooth impression, it is nonetheless only a guess based on the presumption that the face varies continuously and smoothly in each triangular patch delineated by the three fiducial points. Since transformation (morphing) techniques in general rely on the assumption that shape changes smoothly in regions that are not explicitly sampled by delineation points, it is important to capture sufficiently-many sample points. In this regard, three-dimensional sampling techniques are advantageous in creating far more measurements than the earlier two-dimensional techniques, and far more indeed than the original manual callipers-based measurements.

*Three-Dimensional Surface Measurements:* Digital scanning and photogrammetry of an individual face creates a large dataset of three-dimensional points. A polygonal mesh of vertices is computed and, with photogrammetric data, that mesh is texture mapped with the skin colour and texture of the individual's face. By subsequently placing those individual meshes in spatial registration and resampling them it is possible to then derive a standardised mesh of vertices, and in the three-dimensional counterpart to two-dimensional image averaging, a three-dimensional average can be computed along with a computed average of its skin shadings and texture.

The three-dimensional meshes derived by photogrammetry provide facial coordinates that are both precise (Khambay et al., 2008) and dense, with as many as 70,000 vertices (Blanz and Vetter, 1999; Blanz et al., 2007). The resulting mesh can be used as a foundation for subsequent transformation, as a ‘morphable model’ with which to synthesise faces, blends, and caricatures (Blanz and Vetter, 1999; Zhang and Badler, 2006, 2007; Blanz et al., 2007), and as a foundation for face recognition (Section 2.4.1).

## 2.3 Describing Faces

Section 2.2 concerned facial measurement; the following shifts focus to facial description. As stated earlier, measurements are not descriptions. A discrete anthropometric measurement such as the width of the mouth *ch-ch* provides information, of course, but does little to *describe* the mouth — and is less descriptive than saying a given mouth is wide or narrow. A measurement must be interpreted to contribute to a description. Measurements are more akin to depictions than descriptions, in fact: with sufficiently-many surface coordinates measured across a face, the identity of an individual may be detected or that face may be replicated by stereo lithography. Measurements can also lay the foundation for analysing how faces vary spatially, and by understanding the principal components of their spatial variation over some sufficiently homogeneous population, measurements of an individual face along those dimensions can provide the basis for automatically identifying a specific face. As discussed, some facial measurements can also provide insight into statistically-reliable ethnic differences, of course, but a different process of interpretation is required to derive descriptions from such measurements — specifically, descriptions that are useful for humans.

### 2.3.1 Descriptions versus Depictions

Consider the complementary roles of depiction and description in appreciating the human face. A face can be appreciated through the visual exploration of a depiction (a sketch, photograph, etc.), but that process is often more effective and focused when guided by a supplemental descriptive narrative. Especially when comparing two faces, the characteristic features of each must be scrutinised and compared in search for similarities

and differences. The observer often will shift visual attention alternately between the two objects under consideration, and it may take a different type of concentration to notice the *shared* properties, i.e., to attend to what is not changing as one's gaze shifts. Comparison by repeatedly shifting visual attention between two faces is not easy, however, even for experts, hence it is commonplace to add a narrative to guide this visual exploration.

The practice of providing a written description to direct the visual comparison can itself become burdensome to both the author and the subsequent reader. The writer must balance between laboriously reducing shape to an exhaustive enumeration of seemingly objective statements, and attempting to convey the less tangible shape differences by an essay of subjective impressions and metaphorical appeals. The reader's task becomes increasingly cognitive and deliberative, and decreasingly perceptual and intuitive. Fortunately, if faces are described with a vocabulary that closely matches the primitives by which faces are visually perceived, the description is more efficiently understood and can more efficiently lead to insight as the observer considers the depiction of the face.

### 2.3.2 *Categorical Descriptions*

While the usual approach towards studying the human face is quantitative, the conclusions of an anthropometric study is often expressed in terms of the relationships that were uncovered. As discussed earlier (Section 2.2.1), while absolute measurements provide some statistical insight into variations in facial dimensions across ethnicities, the distributions are unfortunately so broad (due to individual variation) that few absolute measurements can be used to statistically discriminate between ethnotypes (Table 2.2). More frequently, faces are categorised on the basis of their proportions. Indices (ratios of absolute measurements) are assigned names based on empirically-decided thresholds (Section 2.2.2). Naming is an important aspect of describing, hence a typical Chinese skull would be described as *mesaticephalic* (having a skull of intermediate length and width, i.e. neither markedly *brachycephalic* nor *dolichocephalic*), *megaseme* (having a relatively large orbital index), *mesorrhine* (having a nose of moderate width relative to

length) and *mesognathous* (having jaws of moderate size and slight projection), as mentioned earlier. Despite the formality of such seemingly concrete terms, the description relies on arbitrary thresholds to partition the anthropometric indices into intervals. The large statistical variability of the human face within each ethnicity precludes making finer descriptions simply by increasing the set of such labels or categories. Moreover, the original anthropometric indices and neoclassical canons were defined relative to EUR measurements (e.g., Gray, 1918). As discussed (Section 2.2.2), recent studies show that the specific facial proportions based on EUR do not necessarily apply to other ethnotypes (Dawei et al., 1997; Farkas, Forrest, and Litsas, 2000; Choe et al., 2004; Choe et al., 2006). There would therefore be no universal set of categories with which to describe facial proportions, except in the broadest of terms (and thus with limited utility). Similarly, for purposes of identifying individuals, an atlas of features and their characteristics has been developed (Ritz-Timme et al, 2011) that categorises 45 facial traits. The facial regions identified by these traits are specific and localised, but assigned coarse categorical distinctions (e.g., nose bridge breadth would be classified as narrow, average, or broad, and chin shape would be round, square, or pointed).

Stepping back from the specific drawbacks of proposing a set of categories based on certain specific empirical thresholds, consider whether computing the quotient of a pair of measured dimensions then selecting a name based on where that quotient falls relative to set of thresholds constitutes providing a descriptor. By labelling a nose that has a measured nasal index of 48.5 as *mesorrhine* gains a name but in fact loses precision. The label is equivalent to a very imprecise, coarsely quantified numerical measurement. And ultimately, measurements are still *numbers*, not *descriptions*. The act of labelling or categorising, such as labelling a nose based on its overall proportions, is nonetheless regarded as adding to its description. It may actually be unfortunate that a categorical term is so evocative, because it appears to provide more concrete and more specific information that in fact it carries.

Before considering the utility of categorical descriptors for describing the faces of different ethnicities, consider a much simpler case: describing four different coffee cup styles (Figure 2.4).



*Figure 2.4. These four styles of coffee cup have obvious shape differences that are apparent in a glance. (They will be termed, from left to right, Styles A-D.) Careful scrutiny is required, however, to find their specific differences, which involves shifting the eyes repeatedly among them, to explore their shapes in order to find places where they differ, and where they share common features.*

Each cup has either a profile that is vertical, tapered, convex, or reflex, a rim that is either large, medium, or small, and so forth. Just as some ethnicities share common facial properties, some cup styles share properties (styles B and C share large rims and medium bases, for example, as indicated by their shared colour in Table 2.4). While this is all neatly tabulated in Table 2.4, that discrete vocabulary captures about as much about coffee cup shape as the terms *megaseme* and *mesognathous* capture about face shape. The illustrations in Figure 2.4 add considerably to the interpretation of Table 2.4. The meaning of the attributes such as “elongate handle” is appreciated at a glance at Figure

*Table 2.4. Styles A-D refer to the styles of the four cups in Figure 2.4, from left to right. Four features are described, each with multiple possible modifiers. Two styles may be compared by examining their respective columns. Styles B and C share similar rims and bases, but differ in profile and handle, for example.*

Feature	Style A	Style B	Style C	Style D
Profile	vertical	tapered	convex	reflex
Rim	medium	large	large	small
Handle	round	elongate	round	elongate
Base	large	medium	medium	small



2.4, for example. Even if one is familiar with the terminology used in the table, without the accompanying visual reference, one can still conclude from the matrix that one cup is tapered with a large rim, medium base, and an elongate handle, but not appreciate the appearance of that combination as a whole.

The visual appreciation of differences and similarities involves directed search and focussed attention. Comparison of different alternatives is ubiquitous across all forms of art and science, such as in typeface design (Figure 2.5). Typefaces such as Gill Sans and Futura would be described individually in terms of a technical lexicon of line weight, stroke width, serifs, descenders, the bias in round letters, and so forth (e.g., Tracy, 1986). Even with these technical descriptors, the two type faces are better appreciated with visual reference to examples of the two fonts side by side.



*Figure 2.5. Typeface styles such as Gill Sans (left) and Futura (right), are conventionally compared by direct scrutiny, alternately shifting visual attention between adjacent examples.*

Physical anthropologists and archeologists likewise utilise a specialised vocabulary for the description of physical artefacts such as different styles of Buddha statues (see Figure 2.6). Phrases such as the “hooked or hero’s nose”, a “chin like a mango stone”, “small snail-like coils of hair” and a “serene expression” refer to stylistic features that are iconographic of different styles of Buddha statue (Rowland, 1963; Diskul, 1991; Fisher, 1993; Stratton and Scott, 2004). Such a vocabulary is incomplete, and the scholars then rely on photographs to complete the description of given style. Buddha styles are then described as some combination of features, some present, some absent. For example, a “nose like a parrot’s beak” is characteristic of the Sukhothai style (1238-1438 CE), but not the Dvaravati (7th-11th century CE) or U-Thong (12th–15th century CE) styles



Figure 2.6. These Buddha statues came from different geographic regions and cultures, and demonstrate distinctive styles. From left to right are examples from Pala, India, Sri Lanka, Pagan, Lan Na, and Dvaravati. Distinctive properties can be tabulated — see Table 2.5.

(Diskul, 1991). The descriptive terminology and their underlying distinctions derive from the literature that studies these artefacts. Closer examination of the feature description in Table 2.5 reveals that these terms hid arbitrary categorical decisions. On what criterion is a forehead regarded as tall versus short, for example? The ad hoc decisions in anthropology required to distinguish a small versus a wide mouth or a high

Table 2.5. A matrix revealing similarities and differences in facial features of Buddha statues from (Wisetchat, 2011, table 3.1).

Feature	Feature Description	Pala	Sri Lanka	Pagan	Dvara-vati	Lopburi	Lan Na	Sukho thai 1	Sukho thai 2	Sukho thai 3	Sukho thai 4	U-Thong	Ayut thaya
Finial	flame/lotus blossom vs. none												
Hair	large hair coils vs. small												
	widow's peak vs. smooth hairline												
Face	elongated vs. broad												
	tall forehead vs. short												
	rounded jaw vs. square												
	lean vs. plump												
Eyebrows	sculpted vs. realistic												
	high arch vs. low												
	continue down nose vs. across												
Ear	sculpted/curled earlobes vs. straight												
Eye	eyelids re-curved vs. straight												
	upper eyelid sharply creased vs. smooth												
	downcast (or closed) vs. direct gaze												
Nose	hooked (hero's) vs. straight profile												
	narrow ridge vs. broad (rounded)												
Mouth	smiling vs. neutral												
	small vs. wide mouth												
	thin vs. full lips												
	outlined (sculpted) vs. realistic												
Chin	creased (outlined) vs. smooth												
	cleft (dimpled) vs. smooth												

Color Code				
Frequency	rarely	seldom	frequently	usually
Magnitude	low	slight	medium	high

versus low arch to the eyebrows of a Buddha statue are similar to those in anthropometry for making categorical distinctions in nose proportion or jaw protrusion in humans faces. A matrix that tabulates the presence/absence of distinctive features of a Buddha style (Table 2.5) helps make explicit their style similarities and differences (Wisetchat, 2011, 2013a). The Lan Na and Sukhothai groups (whose table columns have predominantly dark cells — see key), are clearly distinguished from earlier Buddha styles such as Pala and Dvaravati. The presence or absence of these diagnostic features can also be used to perform a cladistic analysis of shared derived features among Buddha statues (Marwick, 2012).

*Discrete Descriptions of Ethnic Variation:* No sufficiently-rich descriptive written vocabulary has yet been proposed to allow one to adequately describe the average face of a given ethnicity purely by words. However, an important study by Joumana Medlej (2012) compiles sketches of the faces of different ethnicities (Figure 2.7) along with supplemental textual annotations.



Figure 2.7. Sketches of different ethnicities (from left to right): Turkic, Margid Amerind, Nordic, Pacific Amerind, Baltic, South Chinese, and Japanese by Joumana Medlej. From (Medlej, 2012).

The descriptive terminology is intuitive and effective in helping guide the viewer to appreciate the salient features in the illustrations. For example, northern European facial features are annotated as having a “long, narrow face”, “thin cheeks”, a “straight, thin, steep-sided nose”, and so forth, while a classic Mongoloid face is distinguished by a “low nose root”, a “long nose bridge”, “long face”, “small eyes”, “prominent cheekbones”, a “wide face tending to square on the bottom”, and a “small mouth” (Medlej, 2012). A Japanese face would share some of these typically Asian characteristics but be distinguished (according to this study) by annotations including: a “‘pointy’ profile,

thrusting forward”, “... the surface between eyes and eyebrows is not flat but has depth”, a face that is “longer and thinner than other Asian types”, “single eyelids”, a “high, thin small nose, longer in men”, and a “wide mouth” but “thin lips”.

Regardless whether these descriptive phrases are statistically-valid characterisations, they are noteworthy in 1) how they identify specific attributes of specific facial features, such as eyebrow protrusion, cheekbone prominence, and so forth, and 2) how they guide the reader to scrutinise those aspects of the faces in the sketches. The individual written descriptions for a given ethnicity are not readily compared across ethnotypes without further structure, however. A matrix of features could potentially be compiled (as were the distinctive features of Buddha Statues in Table 2.5) to try to organise these separate descriptors, but attempts to describe facial attributes by short phrases such as these results in a loss of precision that severely restricts the utility of such a table. Instead, the descriptive phrases serve better as annotations of sketches that are reduced in complexity to illustrate the essential characteristics of each ethnicity. Then one relies on the shape of the facial contours to gain some insight into those shape qualities. But that raises an important issue: presuming that the above sketches are representative of the corresponding ethnicities, they serve to illustrate distinctive characteristics, even if they are not necessarily dimensionally-accurate representations of the average face of each corresponding ethnicity.

Moreover, in each of the above cases, the depictions are nearly self-sufficient in presenting the corresponding ethnotypes; the supplemental text is helpful in guiding one’s appreciation for the distinguishing features of each type, but, in fact, words are not *necessary* to appreciate the variations they depict. On the other hand, in the absence of the illustration, the supplemental text does not stand alone, neither to convey the shape of each type (of cup, font, statue, or face), nor to convey their differences across type. Likewise, matrices that permit column-by-column comparison of features for alternative types are not readily appreciated without resorting to visual comparisons.

The task of describing an object is different from that of describing the *difference* between that object and another, or that of describing the commonality between that object and another. A matrix such as Table 2.4 or Table 2.5 helps the reader identify similarities and differences, but those similarities and differences remain implicit. Two broad approaches towards representing and visualising differences are described next.

### 2.3.3 Visualising Shape Differences

Consider again the presentations of various shapes of coffee cup, or typeface styles, or styles of Buddha statue, or sketches of various ethnicities (Figures 2.4-2.6). In each case the alternatives are presented simultaneously, arranged from left to right as usual in comparative studies. To appreciate shape differences and similarities, one naturally scrutinises each example individually, shifting gaze from one to the other and back, attending to them serially to build up an understanding of their similarities and their individuality. In Figure 2.4 the depicted alternative cups were identical in size, colour, viewpoint, etc., in order to isolate their shape differences. The shape differences in the Buddha statues in Figure 2.6, however, are not so easily appreciated because the illustrations are of actual artefacts that differ in material, patina, and quality of preservation, and were photographed with different lighting and camera view. As one shifts gaze from one photograph to another, it is more difficult to attend only to the underlying shape differences and to disregard the extraneous aspects presented by these photographs. Shape visualisation is thus facilitated by modelling, so that the presentation can better control for such irrelevant distractions.

*Geometric Morphometrics:* The visual appreciation of shape differences is assisted by explicit depiction using the graphical techniques of geometric morphometrics (the study of form and shape). Rather than leaving shape implicit and up to the observer to discern by observation, shape is mathematically quantified and graphically visualised by tools that isolate shape as a measurable quantity.

Morphometrics distinguishes between related terms: *shape* and *form*. The term shape is used "... to denote the geometric properties of an object that are independent of the object's overall size, position, and orientation, whereas the form of an object comprises both its shape and size" (Bookstein, 1991). In conventional anthropometry, therefore, absolute measurements between landmarks quantify form properties while indices relate to shape. In other words, shape is "... all the geometric information that remains when location, scale, and rotational effects are filtered out from an object" (Kendall, 1977).

D'Arcy Thompson's treatise regarding growth and form (Thompson, 1915, 1917) was one of the earliest, and most influential, methodological contributions towards making geometric shape and shape differences explicit. In a classic demonstration of measuring shape changes, an outline drawing of a fish in side view is overlaid upon a rectangular grid. Distorting the grid distorts the shape of the fish outline lying on that grid. The grid can then be adjusted to morph the outline of one fish to match the shape of another fish. The grid itself is a graphical depiction of the local differences in the two shapes. The distorted grid very elegantly illustrates the shape *differences* between the two fish. This powerful idea has itself morphed into many analytic tools, including the delineation and 'morphing' deformation of two-dimensional images and their three-dimensional morphing extensions (Section 2.2.3).

Morphometrics is concerned with describing shape changes or variation, such as those that occur during growth or evolution, or between related objects. To compare two shapes A and B, there must be a correspondence (or homology) established between points on A and their counterparts on B. Geometric morphometrics, like conventional anthropometry, is therefore based on landmarks, the distinguished points that permit establishing a homology (Bookstein, 1991). The degree of deformation or change from A to B can then be indicated by various means, including graphing the distortion by a rectangular grid or 'thin-plate spline' based on an analogy with a rubber sheet (Dryden and Mardia, 1998), or by so-called 'lollipop diagrams' where vectors are placed at the landmarks indicating the direction and magnitude of their local displacements

(Klingenberg, 2013). Extending the concept of distorting an underlying two-dimensional grid to three dimensions, a scanned three-dimensional surface model can likewise be warped (Wiley et al., 2005; Drake and Klingenberg, 2008).

*Animation:* While static, two-dimensional graphical depictions of morphometric change are best suited for conventional publication, demonstrating the change by watching an animation is often far more effective. The shape change that an object undergoes during some gradual deformation process, or some growth process, is often visualised by stop motion photography. If only given the ‘before’ and ‘after’ shape, however, the intermediate shapes must be interpolated, and to use linear interpolation if there is no other indication of the specific rate of change. Linear interpolation allows a movie to show one object as it transforms into another other. Beyond its obvious appeal of animation for visualisation compared to viewing a static depiction, animation assists in directing visual attention to shape changes across a complex surface, as discussed next.

*Animated Interpolation Directs Visual Attention:* Shape is usually compared by presenting examples side by side, as in Figures 2.4-2.6. In the simplest case, where two alternatives A and B are compared (as in the two typeface examples in Figure 2.5), differences are detected by alternately gazing at A versus B. But instead of statically presenting A and B as separate adjacent illustrations, imagine viewing only one object that is able to change its shape so that it can smoothly transform (or ‘morph’) from shape A into shape B. By keeping the object’s size and position constant while its shape changes, those regions where the shapes A and B differ will necessarily shift or move or distort during the transformation from A to B. The innate ability to shift our attention based on detecting movement can therefore be used to direct our attention to shape differences, as demonstrated in the case of similar styles of Buddha statue (Wisetchat, 2011, 2013a).

Visual attention is unconsciously induced by movement (Helmholtz, 1867, 1962). While attention shifts are most obvious when we notice movement “in the corner of the eye”

and shift our gaze to investigate the cause, attention shifts also occur when directly focussed upon an object and scrutinising it as it changes shape, such as watching a human face for expression changes. Two-dimensional image transformation sequences, where one face morphs into another, e.g., using WebMorph (DeBruine and Tiddeman, 2017) also dramatically draw attention to differences between the two faces.

This automatic redirecting of attention based on motion is useful in comparison of two complex shapes, such as Buddha statues or human faces, using the technique of *animated interpolation* with one model that can change shape to represent different alternatives. As the model morphs from one to another style or shape, our attention is naturally drawn to those shape differences. If two complex shapes A and B differ subtly in *many* aspects, such as the various stylistic differences in Buddha statues or differences in two related ethnotypes, those differences may be best appreciated through scrutiny of an animated interpolation. As a model smoothly interpolates its shape from A to B then back to A, one has the opportunity to notice then scrutinise subtle shape differences, and also to attend to where shapes A and B are similar in shape and thus relatively unchanging. Interpolated animation may better reveal complex shape differences than conventional side-by-side presentations of alternatives, or static morphometric diagrams, and of course more readily visualised than conventional written descriptions.

## **2.4 Face Spaces**

The range of possible variations in the human face is termed a ‘face space’ (e.g., Valentine, 1991). Several of the properties of a face space can be motivated by considering a much simpler space — a space of possible colours. Any colour is specified by a triple of independent variables or degrees of freedom, for example the familiar red, green and blue components of an RGB colour, or the hue, saturation, and value components of an HSV colour. The triple of component values can be regarded as the coordinates of a point in a three-dimensional space. Any specific colour then projects to a point in an RGB or HSV colour space, and similar colours project to nearby points in either colour space. The two colour spaces differ, however, in what information their dimensions make immediately available about a given colour (e.g., the amount of red in



one case, or the saturation in the other). While a colour space is often depicted graphically (e.g., as a ‘colour cube’), the space itself is actually a mathematical abstraction: the set of all possible combinations of the three components. The particular choice of coordinate axes, however, depends upon the application, hence the utility of alternative colour spaces.

A space can likewise be defined to represent faces, where each face would have many more degrees of freedom than the three needed to represent colours, of course. But just as a specific colour maps to a point in a colour space, an individual face (selected from a homogeneous sample population of like ethnicity, sex, age, etc.) would ideally map to a point in ‘face space’. Since the most common application of the face space concept is recognition of familiar individuals, the underlying space will be called an *Identification Face Space* or *IFS*. In contrast, this work will focus on an *Ethnic Face Space* or *EFS*, that would represent different ethnotypes.

#### 2.4.1 A Face Space for Recognising Individuals

The face space concept derives from psychological studies of how individual faces are represented in order that the identity of an individual can be recognised. This has led to algorithms for automated face recognition (Turk and Pentland, 1991; Valentine, 1991; Hancock et al., 1996; Treves 1997; Burton and Vokey, 1998; Busey, 1998; Valentine, 2001; Blanz and Vetter, 2003; Burton et al., 2005; Holub et al., 2007; Valentine et al., 2016). In practice, an individual face might not be expected to map precisely to one point; it has been suggested that a specific individual face maps to a region, or to a manifold, in face space (Valentine, 2001). In modelling the human ability to recognise familiar faces, a face space is also expected to be a ‘perceptual space’, i.e., somehow implemented in our visual system (Shepard, 1962, 1987). The origin of the space would represent the mean value of the population on each such dimension, and it is assumed that faces will form a normal distribution on each dimension, i.e. a multivariate normal distribution in face-space (Burton et al., 1994). A face space is also expected to be a ‘similarity space’, wherein the more typical the face, the closer it will project to the

centre of face space, and the more similar two faces are, the closer they will lie in face space, as measured by their Euclidean distance but “... the similarity metric cannot be determined because the dimensions of the space are not known” (Valentine, 2001; O’Toole et al., 2001).

It has been estimated that 15-22 dimensions would be required to account for our ability to recognise ‘same-race faces’, i.e., individual faces of the same ethnotype (Lewis, 2004). The ability to recognise a familiar individual face (either by visual perception or by automated algorithm) usually presumes a homogeneous population, for which a mean or average face has been determined. Any individual face would be described relative to that mean, and the greater the relative difference, the less typical is that face. The expectation that there is an average or typical face for a given population leads to testing aesthetic notions of ‘ideal form’ and beauty (Langlois and Regimen, 1990; Valenzano et al., 2005; Bashour, 2006; Atiyeh and Hayek, 2008), the role of symmetry (Perrett et al., 1994; Scott et al., 2010) and sex (Perrett et al., 1998; Holzleitner et al., 2014) on facial attractiveness. This central tendency presumption — that the sample population from which an individual is to be recognised has an average face — is a requirement of Principal Components Analysis (PCA), which defines the axes of an IFS for automated facial recognition. It does not apply across ethnotypes, however (Section 2.1.3).

Principal components analysis is a process of fitting an n-dimensional ellipsoid to a collection of n-dimensional data points. The axes of the ellipsoid correspond to the principal components on which the data varies. The first principal component is the longest axis of the ellipsoid, i.e., the dimension on which the data have the greatest variance. The second and subsequent principal components are all mutually orthogonal and account for decreasing amounts of the variance in the data). Strauss (2010) summarises that PCA

“... is used to redistribute the total variance among a set of data points onto a set of mutually orthogonal axes (i.e., at right angles to one another) that merely redescribe the patterns of variation among the data. The new axes are the

principal components, which are statistically independent of one another and so can be examined one at a time.”

In performing PCA, the first step is to centre the dataset at the origin by computing the mean value along each dimension in the original data, which is then subtracted from each data point to make all data relative to the mean. Typical faces are imagined to be cluster tightly near the origin, while distinctive faces would be outliers — but caution is needed when applying intuitions based on two-dimensional clusters to spaces of many more dimensions (Burton and Vokey, 1998). A homogeneous (or so-called ‘same race’) population of faces may well be a multivariate normal distribution (Button et al., 1994), however, if multiple distinct ethnotypes are combined into one face space there would be multiple clusters, not just a single, unimodal distribution, of course.

#### 2.4.2 A Face Space for Describing and Comparing Faces

While an IFS captures individual facial variations *within* an ethnotype, an EFS would describing differences *across* facial ethnotypes. Since there are fewer distinct ethnotypes defined than individuals within a given ethnotype, and EFS would be expected to require fewer DOF than an IFS. The faces of different ethnicities would project to different points in an EFS, but neither absolute nor relative anthropometric measurements could likely serve directly as the EFS axes due to their statistical inability to distinguish ethnotypes, as discussed earlier (Section 2.2, Tables 2.2 and 2.3). The intuition that different facial types *should* be able to be mapped into some multidimensional space is nonetheless very appealing, even if anthropometric measurements would not serve as the coordinates.

Despite countless observations regarding differences in the facial characteristics of Mongolian, Han Chinese, Japanese, and Korean (all within the broad EAS ethnotype), no account has been proposed regarding the dimensionality of an EFS that would reliably map these ethnicities. As with colour spaces, the particular choice of attributes to use as EFS dimensions would depend upon the application (e.g., recognition versus

description). The goal here is to create an EFS to facilitate describing and visualising ethnic variation, not an EFS for automated ethnicity recognition. The sorts of EFS dimensions that would be useful in describing faces should be local, individually comprehensible, and adequate for creating quantitative measurements of ethnic difference, which further suggests that they will not correspond to the principal components derived by some holistic statistical analysis.

Unlike an IFS for which the average face of a given (homogeneous) population lies at the origin and variations are normally distributed about this mean, there is no ‘average face’ to represent the mean across *all* human populations (Section 2.1.3), nor is there reason to presume that facial attributes are normally distributed about such a mean. Instead, cluster analysis of the human genome (Section 2.1.2) shows a pattern of divergence of ethnotypes into distinct clusters that roughly correlate with the major geographic partitions of the globe. Furthermore, there is considerable variation or scatter within each cluster creating overlapping, diffuse boundaries across clusters.

It is only in the broadest of terms that it is useful to regard there being an average EAS face (or an average Japanese face for that matter, despite their relatively greater homogeneity compared to EAS as a whole). The human face cannot be described as a set of variations on one single mean computed across all ethnicities. In particular, since PCA requires a multivariate normal distribution, a principal components analysis of all ethnicities pooled into one dataset would not provide us the orthogonal dimensions or axes of an EFS for ethnicity recognition. Strauss (2010) puts it succinctly: “It’s not uncommon for researchers to use principal component analysis (PCA) to attempt to discriminate groups of individuals. However, PCA is inherently a single-group procedure and is not guaranteed to find group differences even if they exist.” And furthermore, even if the shape of the human face were a multivariate normal distribution, the holistic principal components would not be useful dimensions for *describing* facial variation; the axes need to be semantically meaningful and relate to local features of the human face.

### 2.4.3 Navigating Face Space

Visualising four-dimensional data (just one more dimension than can be plotted directly in a three-dimensional graph) is quite challenging, let alone imagining a face space with many more dimensions. To get around our cognitive limitations, high-dimensional data is often projected from a higher to lower dimensional space (see survey of techniques in Keim, 2002). For example, three dimensions of higher-dimensional data might be plotted in three-dimensional space by a ‘scatterplot’, and then other dimensions added by creating a matrix of scatterplots (e.g., Elmqvist et al., 2008).

Instead of attempting to directly visualise high-dimensional data, it is commonplace to let the space remain largely hidden and to only visualise a specific *point* in that space, the ‘current position’, plus provide the ability to take steps along various dimensions away from that current point. This is a standard method in colour picking (Jalal et al., 2015) where a patch of the ‘current’ colour is displayed and, through a user interface, one can navigate away from that colour along various trajectories (including simply following one axis, such as red in an RGB space or saturation in an HSV space), at each step replacing the current colour patch. To navigate in a colour space, user interfaces might allow directly picking a specific point in a colour gamut display, or choosing one colour sample from an array of samples, or it might allow continuous adjustments of the three colour components by means of slider widgets. Through multiple steps of selection, one can end up far from the original starting point.

Analogously to traversing a colour space in search of a specific colour, a face space might be traversed to search for a specific individual face. In essence, the task is to repetitively replace a given sample face with one that is (at least somewhat) closer to the target, by using a *gradient ascent* technique (Chen, 2003; Chen and Fels, 2004). In gradient ascent, multiple variants on the current choice are arrayed around the current ‘best choice’ then the user selects that variant that is closer to the target of the search than the current best choice, which is then updated and the process repeats until a suitably

close match is found. This sort of algorithm is heuristic at best, as one can become lost or follow various false trails with little sense for where one is or where one is going. Since the search procedure involves choosing an alternative (colour or face), the user of the gradient ascent search need not have any knowledge of the underlying space. In fact, principal components analysis can be used to generate a hidden space of alternative faces (Blanz and Vetter, 1999; Chen, 2003; Zhang and Badler, 2006).

Rather than burrow about in a very large dimensional space, another option is to greatly simplify the space. Two alternatives can be explored. First is the option of isolating a subspace; reducing the overall complexity of face space by attending to only one facial region at a time. For example, the EFS for an entire face could be composed of separate subspaces, such as the space of ethnic variation in noses, or eyes, and so forth. Another alternative is to create a set of exemplars and to then only allow navigation in terms of those exemplars. In colour space, that would be equivalent to permitting linear combinations of only a limited set of colours. In a face space, that would be equivalent to providing access to only a small sampling of ethnotype examples out of the huge space of possible facial combinations. That very reduced face space would then permit navigating only by linear combination of those faces.

### 3. Methodology

This study proposes a scheme for facial shape description and visualisation. The research involves the development and evaluation of a three-dimensional ‘virtual face’ the shape of which is specified through a descriptive ‘language’ and visualised graphically. The virtual face is represented visually by a parametric morphable three-dimensional polygonal mesh controlled by an interactive user interface, called the *Ethnicity Modeller*, or *EM* (Chapter 6). The utility of the entire parametric modelling scheme hinges on its ability to serve as both a descriptive ‘language’, allowing comparison of facial features of different ethnotypes, and a visualisation tool for seeing those differences (not simply quantifying and tabulating them).

The *spiral model*, a process used in software development (Boehm, 1988), was adopted for the development of the *EM* proof-of-concept implementation. The spiral model provides a means to iteratively refine a project that has many unknowns and risks at the outset. The spiral model conventionally starts with an initial concept of the desired requirements and operations, from which a scaled-down prototype is built and evaluated to gain insight into development and operation risks. Following analysis and planning based on the first prototype. The first prototype often reveals concept weaknesses, missing requirements, and technical risks. A second prototype is then developed that addresses issues that were revealed, followed by more detailed validation and testing, and further refinement as the spiral continues.

While the development of the *Ethnicity Modeller* is intended as a only a ‘proof-of-concept’ demonstration, rather than a deliverable ‘turn-key application’, the spiral model was adopted because at the outset of this research there were many unknowns and thus many risks to be evaluated, alternative techniques to be prototyped and compared, and timing was critical, and perhaps most importantly, the development required an initial round of experimentation to drive subsequent refinement of the ideas, all factors that suggest using the spiral model in deciding on how to choose the underlying techniques on which to build the *Ethnicity Modeller*.

### 3.1 Creating a Proof-of-Concept Implementation

The *model* of a specific face of an individual or the representative averaged face of an ethnotype, is represented by a set of attribute-value pairs which can be stored as a human-readable file (in JSON format, see Section 6.5.1). This model file is then read by an *Ethnicity Modeller* to visualise the data contained in the model. For the Ethnicity Modeller to be useful for this study, it provides the following six functional capabilities:

1. The surface geometry can ‘morph’ (Section 6.3) under parametric control to closely approximate the surface of a representative sample face of a given ethnotype (either that of an individual or the averaged face of a population sample).
2. Objective criteria can measure the quality of fit between model and a given sample face (Section 6.5.3, Section 7.1.1, and Appendix A3.1).
3. The set of control parameters governing the surface geometry of the model correspond one-to-one with facial attributes. The modelled facial attributes would then be adjusted only through these parameters.
4. Attributes are represented as interpolates (or blends) between three-dimensional exemplars. The value associated with an attribute is an interpolation coefficient specifying some fractional extent towards the extreme.
5. The parameter values (each representing a normalised attribute value) are comparable across models, as they are all interpolates between the same extremes. Differences and similarities in attribute values between ethnotype models quantify the facial differences and similarities of the corresponding ethnotypes, providing both descriptive and comparative utility.
6. Ethnic variation can be visually appreciated by observing the changes in one model as it morphs from one to the other is core to this approach (Section 7.2.3). These changes can be appreciated through manipulations performed at different levels of complexity: across the whole face, by selected facial region, by isolated feature, and by isolated attribute of a given feature.



The face representation and visualisation scheme was evaluated (Section 7.1) and shown capable of capturing a broad range of ethnic variation by adjustment of a set of facial attributes. The reference material for validating the accuracy of a model will be three-dimensional samples of individuals or computed averages from a sample population of a given ethnicity. With the two surfaces present in the same virtual space and carefully aligned, the quality of the fit between the model and the sample was measured by examining the spatial separation (in millimetres) between the two surfaces.

Three-dimensional surface modelling will provide a basis for visually appreciating the shape and proportions of the human face for various ethnicities (Section 7.2). Moreover, the development of a common parametric face model, once calibrated, will permit modelling different ethnicities and, through interpolation, blending between ethnicities. For instance, if the populations of two geographically-separated areas A and B have distinctly different faces, is the average face of people living between A and B an interpolation of those two types? Exploration of ethnic differences in the appearance of the face will be achieved by manipulating the values of the model's attributes.

## **3.2 Synopsis of the Methodology**

### ***3.2.1 Defining Facial Attributes***

The first step in describing the face was to decide upon the descriptors (Chapter 4). The conventional set of soft-tissue landmarks provided a framework that places this work in common with traditional anthropometry (Section 2.2.1). Lengths measured between landmarks (and proportions computed as quotients of such lengths) are of only limited use in characterising faces (Section 2.3), however, consequently the majority of attributes defined in this study concern the properties of extended features, not point-like locations. The methodology by which these attributes were identified is as follows.

1. *Collection of a photographic data set.* Since individual variation within an ethnicity obscures the facial characteristics common to that ethnicity, a representative sample set of 20 or so individuals for each of several broad ethnicities were photographed, then delineated and image averaged in WebMorph

(DeBruine and Tiddeman, 2017) to create averaged images of those ethnotypes; see Section 2.2.3. Photographs of both front and side views were taken of roughly 100 individuals. These uncompensated volunteer participants understood that their images would only be used anonymously, only applied to this immediate research project, and that they could withdraw their participation. These images were 2000x2000 pixels, providing sufficient resolution to permit cropping to isolate an individual facial region (cranium, eyes, jaw, midface, mouth, and nose).

2. *Feature analysis per facial region.* For each selected pair of ethnotypes (e.g., EUR and EAS), closeups of each facial region were cropped out and used to as the beginning and end frames of a continuous morph or blend movie of that region between pairs of ethnicities, using WebMorph (Wisetchat, 2018). Starting with the eye region (Section 5.1), a total of 17 attributes (across 13 distinctive features) were readily observed to differ with ethnicity, named, and compiled (Table 5.1). The same process was repeated to compile salient attributes for the other facial regions, purely on observation of the two-dimensional morph movies. A total of 77 attributes were eventually identified.

### 3.2.2 *Modeller Development*

All models are based on a common framework, which is called a *modeller*. A modeller consists of three elements: a polygonal base mesh, a set of target meshes, and a deformer (Section 4.10). A specific *model* is created only through the deformation of the base mesh, and consists of the set of deformer weights. Development of the proof-of-concept Ethnicity Modeller, the *EM*, demonstrated the framework of a modeller for visualisation, plus a set of models for various ethnotypes (Chapter 6). The modeller was developed using the Spiral Model (Boehm, 1988) process. In the first iteration of the spiral, the requirements for the deformer were determined, alternative experimental deformer implementation techniques were implemented and evaluated, and it was decided that the EM would be based on blendshape deformers (Section 6.3). Development then proceeded through an implementation-validation-plan cycle, as follows:

1. *Iterative development of a polygonal face mesh.* A polygonal mesh was developed with sufficient topology to support the modelling of facial attributes. Adjacent facial regions (such as the nose and mouth) necessarily share vertices along their common borders, causing challenges for the design of separable facial features. The creation of the base mesh involved many iterations of refinement as facial attributes are identified and added (step 2).
2. *Implementation of facial attributes.* The facial attributes (Section 3.2.1) were then implemented by basis shapes (copies of the base mesh) that represent extremes of each given attribute in isolation (and blended into the base mesh according to an interpolation value by a blendshape deformer (Sections 6.3 and 6.4). The process of creating a suite of basis shapes was laborious and required iterative modification of the polygonal base mesh in order to provide adequate mesh topology and detail. While the discrete facial attributes were compiled by examination of two-dimensional imagery, the creation of three-dimensional representations of their extrema soon required shifting to better, three-dimensional (stereo-photogrammetric) face data, especially to understand the subtle three-dimensional shape of features such as sulci, and fossae not apparent in photographs (Section 6.5.3).
3. *Validation.* During the recurring process of validating the representation of attributes by pairs of blendshape targets, the *EM* was used to model various samples (both averaged data for a given ethnotype and data from individuals within an ethnotype) in order to test whether their features could be captured in the *EM*. A successful model would be one in which the polygonal mesh and the empirically-provided data mesh differed less than the variance expected due to individual differences, typically a few millimetres (Section 6.5.3).
4. *Plan.* After a period of implementation and validation, revisions were planned to improve the topography of individuals basis shapes and sometimes the topology of base mesh itself (which then needed to be ‘baked’ to the all target basis shapes through a laborious process). The planning concerned attempting to isolate the effects of one attribute upon another, for while attribut-

es are ideally orthogonal, their implementation can have unexpected interactions given the finite nature of the underlying polygonal mesh. Combinations of attributes must be examined to detect and resolve interactions. To achieve this requires incremental adjustment of the associated target meshes of that attribute and, if necessary, modification to the base topology. Often the range of variation required of the attributes was greater than originally expected based merely on the two-dimensional images. Also, in the process of attempting to create adequately close matches with such data, it became clear that additional attributes needed to be added, with the ripple effect of changes that would entail.

### 3.2.3 Evaluation

Two fundamental aspects of the *EM* were evaluated: first, its ability to represent faces of differing ethnicity with sufficient precision and accuracy, and second, the usability of its user interface.

The representational ability of the *EM* had already been rather exhaustively tested in the process of implementing and perfecting the attribute basis shapes during the earlier development stage. Creation of a given model involves adjustment of the deformer parameters until the base mesh deviates from the sample mesh by less than the sample uncertainty. First the sample mesh is aligned with the base mesh such that the corneas of the sample mesh are precisely superimposed on those of the modeller. The deformers of each facial region are then adjusted until the base mesh closely approximates the sample mesh (the surfaces overlap and differ by less than the given sample uncertainty). The modelling of individuals of various ethnicities was unexpectedly useful, for frequently one or another facial attribute would require adjustment far from the usual value for that ethnicity. Efforts to modify the basis shapes to accommodate increased range often improved the representation of that attribute over its entire range. As a result, the attributes could model not only ‘average’ faces of various ethnicities, but also faces that were atypical (e.g., an especially protruding gnathion, or an unusually narrow gonion).

Once the *EM* was regarded as ready for evaluation, the capabilities of the representation were then tested for two specific criteria: precision and accuracy. Precision was measured by a repeated-measure study (Section 7.1.1). Each of two ethnotypes were modelled repeatedly 10 times, after which the statistical distribution of attribute values across repeated trials was examined. Accuracy was measured by a related task, in which a select set of dimensional facial attributes were calibrated in millimetres (Section 7.1.2), so that these *EM* measurements could be compared with the corresponding, established, anthropometric measurements for that ethnicity.

Regarding the user interface, the *EM* was presented to a set of computer-literate users with a variety of skill level in three-dimensional modelling from non-expert to expert. The ‘Thinking Aloud’ methodology (Lewis, 1994) was used, an interview technique in which typically four to five individuals were guided through usage of the *EM*, and encouraged to describe their experiences as they attempted to achieve a predetermined set of tasks.

While the *EM* was intended only for expert use in development and validation of the current scheme for representing and visualising ethnic variation, it was nonetheless important to evaluate how well the proposed attributes capture facial features. Were they ‘semantic’ and intuitive, as hoped? The five participants in this study understood that their participation was voluntary, and that their participation and comments would be anonymised.

This evaluation process through interview sought specific feedback on how well the proposed set of facial attributes could be treated as ‘semantic’ face space dimensions, how readily they could adjust the morphable face mesh to approximate a specific nose, and whether the attributes correspond to facial features as seen by these participants.

### 3.2.4 Model Manipulation and Ethnicity Visualisation

A set of tools will then be created to permit exploration of ethnic differences. As discussed in Section 2.3.3, a primary tool for visualising ethnic differences will be through animated interpolation. The common idea here will be to create a continuous morph between two faces. These include:

1. *A-B Interpolation.* Two models, A and B, can be compared by continuous interpolation that results in the base mesh appearing to smoothly alternate between A and B. The interpolation may be on the entire face, or by individual facial regions. This tool will permit study of interpolated facial types as blends of two ethnotypes.
2. *Multi-type Interpolation.* A more complex interpolation can be created by multiple models given geographic locations. This will constitute one of the core Ethnicity Face Space (EFS) interpolations. EFS navigation would mean visualising the human face as it changes *continuously* from resembling one ethnicity to another, through *intermediate forms*. Navigation might be based on underlying geographic locations for the different ethnic centres, or along reduced degrees of freedom.
3. *Extrapolation (Caricature).* As tool for ethnicity visualisation, a given model can be exaggerated parametrically to enhance those attributes associated with that model.

## 3.3 Working Assumptions

This study is based on the following assumptions:

1. *Describing by parametric modelling:* It is assumed that the facial shape characteristic of an ethnotype can be captured sufficiently by a 3D parametric modeller that implements a set of discrete *shape attributes*, each governed by an associated attribute value or parameter.
2. *Common attribute set:* It is assumed that a range ethnotypes can be modelled based on a common set of attributes by choosing different combinations of

their *values*. The resulting *model* of a given ethnotype is represented (or encoded) by that set of attribute *values*.

3. *Separable, independent regions*: It is assumed that the geometry of the whole face can be subdivided into facial regions. The description of a whole face would then reduce to modelling its regions independently.
4. *Spatial resolution*: It is assumed that the surface shape of any ethnicity can be replicated parametrically to within measurement error and population variance for that ethnicity. This assumption is directly testable (see Section 7.1).
5. *Sample mean plus representative examples*: To study the facial shape of a given ethnotype, it is assumed sufficient to examine the shape of a *sample mean* from that population, plus the shape of representative *individuals* revealing variation about the mean.

## 4. Facial Description

### 4.1 Local Features

The description of the human face and its variations builds upon a formal lexicon of landmarks and shape features that is used consistently across a broad literature. This thesis study adopts this literature and conservatively adds new terminology only as needed, using the established naming conventions. But unlike conventional descriptive vocabularies, the current system offers the ability to both describe facial features and to visualise them, in one common framework.

The face is naturally regarded as a mosaic of spatially separate, abutting, regions (the forehead, nose, eyes, etc.), each of which can be described in isolation, suggesting that the face as a whole can be described as the sum of the descriptions of those regions. Six regions will be defined here: the cranium, eye, nose, mouth, midface, and jaw (the specifics of each area will be introduced later). Fortunately, the complexity of facial shape in each facial region is largely contained entirely within the boundaries of that region, and the face is smoothly faired along the boundary of each region with the neighbouring regions. There is little spatial overlap between regions and neighbouring regions, and no significant curvature features (such as ridges or troughs) cross those boundaries.

Anthropometric studies of ethnic variation often focus on a single facial region in isolation. The features of each region are named, measured, described anatomically, compared across sex, compared across ethnotype, classified, judged relative to standards of attractiveness, modified surgically, and so forth. The eye region, for instance, has a very complex geometry, with many skin folds, creases, sulci, and fossae, and yet all that complexity within the eye region can be delimited neatly by a roughly elliptical border bounded by the nose, the forehead, the cheek, and the temple. The nose can similarly be isolated such that its shape attributes only pertain to that region of the face, with minimal complications between nasal attributes and those in adjacent regions. This modular quality of the human face is well recognised, of course, and has many implications, such as



the human ability to create facial expressions in virtually arbitrary combinations of contributions from different facial regions. This modularity is welcome as it simplifies the design of the model, and those relatively few exceptions from strict modularity (problematic dependencies where the description of an attribute in one region depends upon aspects of another) can be dealt with in a case-by-case manner.

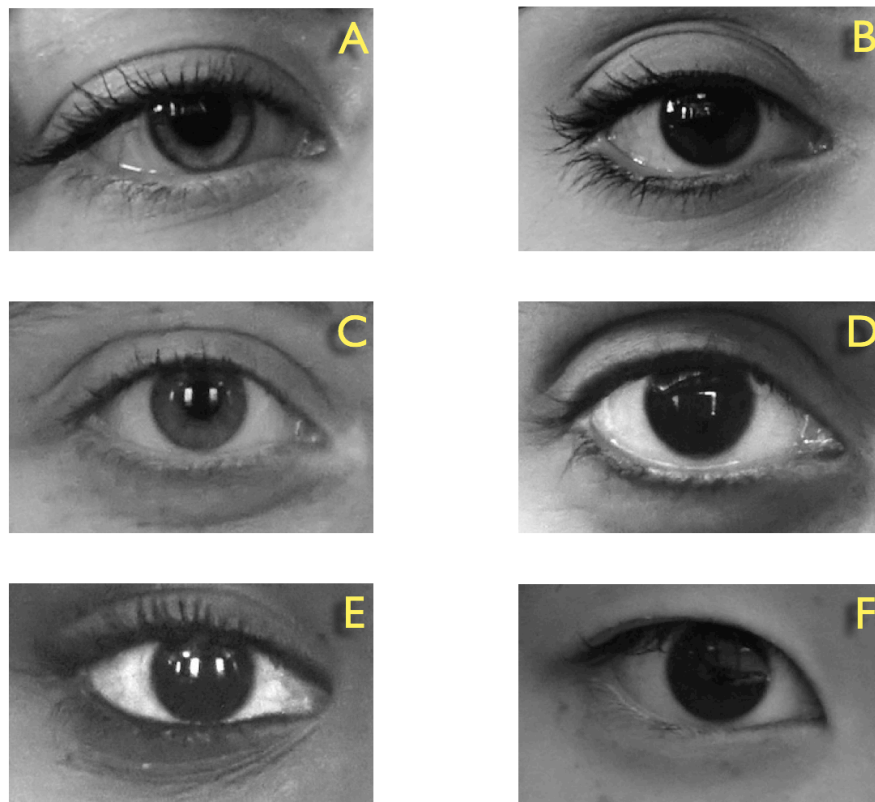
While it is commonplace to describe the face in terms of attributes of features within regions, anthropometric measurements frequently span across regions. That is, a distance is often measured from a landmark in one region to another landmark in another region. For example, the width of the eye fissure is *ex-en* (the separation between the exocanthus and the endocanthus). The overall width of the two eyes is frequently measured from exocanthus to exocanthus (*ex-ex*), as well as from endocanthus to endocanthus (*en-en*), where both distances are measured from a landmark in one eye region across the nose region to its counterpart in the opposite eye. With all three measurements there is in fact redundancy, since  $ex-ex = en-en + 2*(ex-en)$ , as will be noted again below (Section 4.6).

A partitioning of the face into regions that are defined primarily in a frontal view will necessarily correspond to a face seen in lateral view. If the model were strictly two-dimensional, that would clearly be a problem. But this study is building upon three dimensional representations of facial shape. While in a front view the nose might appear completely separable from its neighbouring regions, from lateral view of course it is seen to be a protrusion above the midface. The same nose could be placed upon a relatively flat face or one in which the maxilla projects substantially beyond the plane of the zygoma. In the lateral case, the nose projects further as it rides atop a projecting maxilla. Thus the nose can remain a substantially-independent module, while adding its surface relief beyond that of the underlying face. The overall profile of the face, in lateral view, will be treated as the superposition of many local, substantially independent, attributes.

The goal of this project, recall, is to assist in the visualisation and description of the human face and to appreciate its variation across ethnicities. Fortunately, each facial region

provides some evidence of ethnicity, in the form of distinctive characteristics (of shape, contour curvature, depth of skin folds, and so forth), but at a smaller scale than traditionally measured anthropometrically. To illustrate, consider just the eye region. Figure 4.1 shows the right eye of six females of differing ethnicity. The images are all monochrome to minimise colour clues to ethnicity.

One can appreciate the ethnic differences in Figure 4.1 as reflected by the distinctly different features and details in the six images. Some eyes are deep-set with prominent creases, while others much less so. The shape of the upper and lower margins of the eyelids are also quite distinctive, especially in the inner corner of the eye. These and other variations appear to be strong clues to the ethnicities of the individuals. More specifical-



*Figure 4.1. What ethnicities can one distinguish among these individuals? Clearly A is Western European and F is East Asian. There is also a West Asian (WAS, specifically Saudi), another Western European (WEU), an East African (EAF, i.e. Kenyan), and a South Asian (SAS, i.e., Pakistani). Answer: A = WEU, B = SAS, C = WEU, D = WAS, E = EAF, F = EAS.*

ly, the superior palpebral sulcus (the fold above the upper eyelid) is deep in EUR and AFR, but less apparent, or even absent, in EAS. Perhaps the most distinctive feature of EAS eyes is the epicanthal fold (an extension of the upper eyelid that partly obscures the endocanthion). The epicanthal fold is less frequently found in AFR and is generally absent in EUR. The depth of the superior palpebral sulcus and the prominence of the epicanthal fold are salient characteristics that provide information about ethnicity.

While EAS eyes frequently exhibit an ECF, they may not have an SPS, while the opposite is generally true for EUR. In attempting to create a descriptive scheme of facial features that applies to all faces, how should the absence of a feature be represented? The solution here will be to regard features that are present in some cases and absent in others as having a normalised weight which varies from 0.0 when the feature is absent to 1.0 when that feature is present to some maximum or extreme degree. For instance, regarding the epicanthal fold, that value would be zero, when that feature is utterly absent, increase in value for an increasingly prominent fold, and represented by a weight of 1.0 if the ECF entirely covers the inner corner of the eye to the maximum extent normally observed across ethnotypes.

## 4.2 Attribute-Value Pairs

A description, in general, is compiled out of a language of discrete descriptors. A descriptor applies to some aspect of an object that is summarised in terms of an *attribute-value* pair. An attribute's value may be *quantitative* (i.e., continuous valued) or *qualitative* (assigned a value from a discrete set of categories). Categorical descriptions are less precise, of course, as they partition some continuous value into discrete intervals. The attribute's value can also be either relative or absolute. A value expressed in *absolute quantitative* terms can be interpreted directly, e.g. (weight : 54 kg). Relative descriptions, however, require some reference value such as a mean or nominal value, or a range that is bounded by minimum and maximum values). For example, provided that some property is normally distributed, for which the mean has been computed, a specific attribute value might be expressed in *relative quantitative* terms such as (weight :

$-2.4\sigma$ ) (i.e., 2.4 standard deviations less than the mean), or as a relative fraction such as (weight : 0.8) (i.e., 80% of the way between a pair of minimum and maximum weights).

Dictionaries define a *description* as “a statement or a piece of writing”, “a spoken or written account”, “a statement that gives details”, “a statement or passage”, and so forth. Natural languages build upon discrete vocabularies that, in combination, permit the communication of a vast range of possible utterances. Those discrete descriptors are fundamentally inventions; we choose a vocabulary of nouns and adjectives that are useful for our purposes. In everyday life, the human face is thus described through an invented language of descriptors that refer to facial features, such as a ‘bulbous nose’, a ‘chiseled chin’, a ‘protruding brow’, ‘high cheekbones’, and so forth. Each facial feature has an associated set of modifiers specific to that feature. Some would describe dimensional or size aspects, others shape aspects. In terms of attribute-value pairs, these might be expressed as (nose\_size : large) or (nose\_shape : bulbous). It is doubtful, however, that an adequate description of the face can be devised based on any such set of discrete modifiers applied to discrete facial attributes. While a bulbous nose or protruding brow readily brings to mind examples of individuals with those traits, those visual associations are neither objective nor very specific. What, for example, would be the specific meaning of ‘slightly chiseled’ or ‘very bulbous’? These and other qualifiers make sense only with reference to some standards. Moreover, there would also have to be some constraints on the application of modifiers to a given facial attribute, since clearly some combinations would be mutually exclusive (chiseled and bulbous), and even modifiers that are not mutually contradictory might have unexpected implications when used in combination. Nonetheless, there is something to be salvaged in attempting to describe faces by a discrete vocabulary of facial attributes and their modifiers. In this study, the values associated with facial attributes will be relative and quantitative. A specific set of attributes, and the basis for quantifying their relative values, will constitute a considerable part of this study.

### 4.3 Description by Analogy

A specific face is often described with reference to known exemplars. Attention tends to focus on those properties that make the given face distinctive or remarkable. Those various qualities are often described by comparison with known individuals (“the same nose as ...”), or with those of some population for which that trait is commonplace (“... a Roman nose”), or even an inanimate object that shares some resemblance (“A nose like a ...”). Description by analogy is very natural and commonplace, but in fact it subtly avoids (or at least, defers) the question. When describing some object A as being like some other object B, the latter is not in fact described, and yet we are often satisfied with the answer. Informally, a face is often described with reference to some understood ‘average’ or ‘normal’ face. Often only the distinctive aspects of a face are described, not those which are unremarkable. An average face is thus very difficult to describe by this approach.

Describing some attribute relative to some norm or expected mean value for that attribute presumes that there is some mean or average on which to base this comparison (see Section 2.1.4). Alternatively, the value assigned to an attribute might be described relative to the extremes of variation for that attribute — recall that weight might be represented by the attribute-value pair (`weight : 0.8`). The same concept would apply to an attribute that is not simply a numerical value such as weight, but a shape feature. If that feature varies between two extremes of shape, a given value could be assigned a fraction that would correspond to the fractional interpolation between the two shape extremes. The strategy of describing shape attributes by interpolation between extremes will play an important role in this study.

### 4.4 Orthogonality

While size and shape seem to be independent aspects of a given facial feature or region, in fact, they are not strictly orthogonal, a complication that arises when attempting to be precise about these informal descriptors. To be truly orthogonal, the two descriptors, or qualities, would have to potentially occur in arbitrary combination. Importantly, the no-

tion of orthogonality between descriptors is a conceptual or abstract concept related to non-interference. It does not require that the given combination of attributes actually occur in some individual or population of individuals, but that it *could* occur, in principle. For example, if some noses are to be regarded as ‘chiseled’, it should at least be conceivable that those noses be either large and chiseled or small and chiseled. Some combinations are clearly mutually exclusive, such as a ‘small large nose’ or a ‘bulbous chiseled nose’ as the descriptors refer to opposites. But some other combinations can more subtly lead to contradictions, such as a ‘a tall, narrow, bulbous nose’ where the term ‘bulbous’ connotes “fat, round, or bulging”, i.e., some qualities of size as well as shape. To avoid these issues, the choice of attribute is important. While ‘bulbous’ and ‘chiseled’ are evocative and familiar terms, they will not in fact be attributes used by this study. Reflecting on the adjectives just used above, we see that descriptions are still based on categorical distinctions. While some facial region such as the nose might be measured with callipers between various points on the soft tissue, such measurements are not themselves descriptions. The absolute width of an individual’s nose, measured between the lateral extremes of the nostrils, is but one datum, which may be influenced by that individual’s ethnicity, age, height, sex, as well as specific individual, developmental factors.

As will be discussed, much of the appreciation of the human face and its ethnic variations has derived from the careful anthropometric quantification of dimensions, angles, and proportions. Quantitative measurements are then frequently categorised. A nose, for example, is sometimes categorised as either *leptorrhine*, *mesorrhine*, or *platyrrhine* (see Section 2.2.2) based on the proportion of nose width to nose height. Such discrete, named, categories, however, are arbitrary and will be avoided in the current study. Instead, attributes that vary substantially within a population will be treated as continuous-valued without imposing discrete categorical distinctions on that property.

#### 4.5 The Insufficiency of Measurements as Descriptors

Dimensional descriptors would seem to be a rather direct and uncontroversial path towards describing faces. Given a convention of facial landmarks (chosen primarily on the

grounds of convenience, repeatability, and ease of identification), the careful compilation of masses of measurements taken between pairs of such points, for individuals of various ethnicities, sex, age, and so forth, would seem to offer a path towards eventually fully describing the human face and all its variations. The problem that remains, however, is what to do with such measurements in order to form *descriptions* of the human face and all its variations. Again, our almost universal solution when faced with such data is to define categories, and to seek patterns of the occurrence of those categories in various human populations. Almost immediately a few patterns do emerge in comparing ethno-types. The noses of Africans are generally broader than those of Europeans, for example. In fact that statistical trend has lead to defining categories of noses based on their width, (Section 2.2.2). A few other categories have been adopted that pertain to the protrusion of the jaw, the width of the cranium, the breadth of the cheekbones, and so forth. They *are* truly facial descriptors, and they are based on carefully compiled statistical measurements. But if they were to be regarded as the beginning of a ‘path towards describing faces’, that path would not take us very far. In fact, the few categorical distinctions that emerge from centuries of observation of human facial variation provide only the broadest of guidelines.

#### 4.6 Variance and Measurement Error

In describing the average face of a given ethnicity, it is assumed there *is* some average, i.e., that the notion of the population *mean* is well-defined. To regard there being a single mean AFR face is clearly suspect, since the diverse facial types across such a large continent results in broad (and not necessarily normal) distributions for the various attributes one might wish to examine. Some attributes are more or less shared across AFR (such as nasal width *al-al* which distinguishes AFR from EUR and EAS faces). Likewise, attempting to define a single Asian ethnotype (ASN) is virtually meaningless, as it would combine the very different faces of West Asians (e.g., Arab) and South Asians (e.g., Indians) with East Asians (Mongolians, Chinese and so forth). For this reason, the *East* Asian subregion EAS is chosen for this study, since it comprises the greatest proportion

of the ASN population, and is also fairly homogeneous (primarily composed of Chinese, Japanese and Korean; Section 2.1.3).

When considering the modelling an ethnotype, this study will rely on population samples of various ethnicities, however it will not address their underlying statistical properties. The basic ethnic groups of EUR and EAS will be sampled by averaged three-dimensional stereo-photogrammetric reconstructions (Section 6.5.3).

While it commonplace to publish anthropometric measurements to a precision of a tenth of a millimetre (Farkas et al., 2005), such precision does not consider measurement uncertainty (Taylor, 1997). It is unlikely that anthropometric measurements of soft tissue across a given face can be more precise than a millimetre or so. In addition to measurement error, individual differences for a given anthropometric measurement taken *within* a population are often greater than the differences in mean measurement *across* populations. It is difficult to statistically distinguish between ethnicities where the means differ only by a few millimetres, and yet the standard deviations are several times greater than the difference in means (see Table 2.3). As concluded by Farkas et al. (2005), faces of different ethnicities are more similar dimensionally than they are different, for most dimensions.

The measurements by Farkas et al. (2005) also show numeric inconsistencies that may reflect measurement error. For example, *ex-ex* can either be directly measured by calipers or computed indirectly from *en-en* and *en-ex* measurements. For Japanese males (from their table 22a)  $en-en = 37.5 \pm 3.1$  mm,  $en-ex = 30.7 \pm 2.3$  mm, and  $ex-ex = 103.9 \pm 4.2$  mm. But since  $ex-ex = en-en + 2*(en-ex)$ , that would suggest  $ex-ex = 98.9 \pm 5.4$  mm. The substantial 5 mm difference is probably attributable to measurement uncertainty. Follow-on studies to (Farkas et al., 2005) have underscored the considerable individual variability of anthropometric data (e.g., Fang et al., 2011). In addition to measurement uncertainty and the variance due to individual differences within a presumably homogeneous population, there is an inevitable problem of attempting to measure a broad,



heterogeneous population such as AFR, EAS, or EUR as if it were comprised of a single ethnicity.

Consequently, when it comes to creating a model of a given ethnicity based on some sample population, that model need not match the measurement data closer than the variance in that data. The imprecision in anthropometric data suggests a practical limitation of several mm in any absolute measurement, in general.

#### **4.7 Descriptions of Continuous-Valued Shape Attributes**

The path towards describing faces would appear to take us further if we were to describe facial features in terms of *shape* attributes. While a discrete vocabulary of shape descriptors is often used, particularly for facial features that are relatively unusual or distinctive (an ‘aquiline nose’ or a ‘jutting brow’), these descriptors can be represented in more fundamental terms, such as measures of surface curvature or the relative height of local prominences and hollows. For example, an ‘aquiline nose’ (after the Latin word meaning ‘eagle’) or ‘Roman nose’ is characterised by a distinctive profile. The nose does not form a straight line from base (nasion) to tip (apex), but rather, it rises above that straight line to form a convex, beak-like, curve. The curve of the aquiline nose is but one of a continuum of possible nasal profiles. Since the profile of the nose might be complex due to irregularities along the length of the ridge, a discrete vocabulary of adjectives will not be practical for describing those various profiles, nor would it be very precise. The contour of the nose in profile is of course a consequence of some combination of the shape of the underlying nasal bones and the shape of the dorsal and lateral cartilage and other soft tissues. While the profile shape (whether ‘aquiline’ or ‘pug’) is a visually important characteristic of the nose, that shape need not be explicitly described. Instead, it can be captured indirectly by some combination of descriptors of nose features. The particular choice of what descriptors to make explicit (and what others to remain implicit) will be an important aspect of this research. In describing the face, this study emphasises the importance of three-dimensional descriptors, and de-emphasises the importance of classifying terminology such as ‘aquiline’.

Two points are illustrated by this exercise: 1) the eye region carries information about ethnicity, and 2) that ethnic information is in the form of different combinations of attribute-value pairs of its features. The two eye features mentioned thus far are anatomical: the superior palpebral sulcus (the hollow region above the upper eyelid), and the epicanthal fold (which partly obscures the inner corner of the eye). The superior palpebral sulcus, or SPS, is assigned a single attribute, a measure of its depth, which would range from a negligible hollow or trough above the upper eyelid to a very pronounced depth that causes that region to appear sunken. This range could be normalised from 0.0 to 1.0, and the value of SPS would have some value within that range, corresponding to a specific shape that is interpolated between the shapes corresponding to the two extremes for SPS. Similarly, the epicanthal fold, or ECF, is a feature of the eye that would have an associated weight attribute, where 0.0 corresponds to an eye with no ECF and 1.0 corresponds to the most prominent epicanthal fold expected. Note that neither ECF nor SPS are normalised relative to any presumed population mean, or ‘average face’, as it is intended to apply across all ethnicities. There is no expectation for the average depth of an SPS nor the degree to which an ECF is present. In fact, as has been stressed several times over, the concept of an ‘average’ face (while required for some statistical analyses) is probably not well-defined for many ethnicities, let alone the notion of an average human face. Instead, the approach in this study is to provide the extremes as concrete exemplars (in fact, they will be represented as three-dimensional surfaces), and let each face be assigned a value for each attribute within those extremes. The distribution of attribute values will not be expected to exhibit a central tendency; they are not expected to be distributed about zero. Instead, extremes of variation are provided that are sufficiently broad as to capture the range of variation in faces of different ethnicities (but perhaps exceeded by some individuals).

#### **4.8 Arbitrary but Useful**

This study will build upon an established anatomical and medical lexicon for describing facial features, one that has evolved over centuries of practicing the art of description and measurement of the body. The conventional facial lexicon is comprised of:

1. anatomical locations or ‘landmarks’ some located on the soft tissue of the face and others defined by the underlying osteology,
2. distance measurements between landmarks and proportions (quotients) of measurements, and
3. shape features such as prominences, convexities, shallow concavities, ridges, creases, folds, and troughs.

While this lexicon is certainly not complete in any mathematical sense, it has been sufficient for describing and measuring facial anatomy and physiology. The locations and topographic features are not evenly distributed across the face; rather, they correspond to salient features that are clustered about the eyes, nose, mouth, and the margins of the face. These could well serve as the ‘nouns’ of a facial description system, on which to associate descriptive modifiers. The skeptic might conclude that this path towards describing faces is hopelessly arbitrary, building as it would on an ad hoc feature set. But then, most technical vocabularies are also ad hoc and evolved by need, convenience, and are preserved by convention, rather than by any sense of mathematical sufficiency, self-consistency, or completeness. The landmarks are conventionally used for anthropometric measurements, from which facial dimensions, proportions, and coordinates are extracted. Statistical means are computed to quantify the ‘average face’ for a given population sample, in order to compare across populations or to measure how an individual face may differ from the mean of a given population (e.g., for facial recognition).

The present study finds that the shape of topographic features such as folds and ridges (e.g., the supratarsal fold of the upper eyelid and the supraorbital ridge above the eye) are of greater importance in the *description* of the face, than are *measurements* based on landmarks.

#### **4.9 Exemplar-Based Attributes**

While facial dimensions that are measured between landmarks quantify shape-related features, they do not represent descriptions; those numbers require interpretation. Rarely

can a single facial dimension be interpreted based directly on its absolute value, although it is commonplace to find tabulations that compare various absolute measurements for different population samples. Alternatively, *ratios* of absolute measurements quantify facial proportions, which can also be compiled and compared across population samples. In addition to comparisons, these ratios or ‘indices’ (Section 2.2.2) are used to classify facial proportions. For example, a nose might be deemed *platyrrhine* based on the measured ratio of nose width to height. The result is a description by analogy to an exemplar (nose A, being platyrrhine, resembles some prototypical platyrrhine nose B).

Classifications result in considerable loss of precision. Alternatively a measurement could remain a continuous quantity by employing two exemplars that represent the extremes of that measurement. The nasal index, then, would be converted to a normalised value (nose A is 37% between the most leptorrhine and the most platyrrhine). Precision is no longer lost by quantisation, and a given value for that attribute can be represented by interpolation. The description is not explicit, however, as it is presented with reference to exemplars.

This introduces the strategy used in this study. A set of facial attributes are identified, a range of variation is determined for each, then a given face is described by matching the attribute value for that face in terms of its potential range. The resultant set of attribute-value pairs will together constitute the description of the face. The suggestion, then, is to describe faces by parts:

- 1) A vocabulary of discrete facial attributes will be compiled by first identifying a set of salient facial features. Each feature will be described in terms of its attributes. Some facial features are spatially-extended portions of the face such as the forehead or cheek, others are very localised such as the shape of the corner of the eye, or the depth of a fold of tissue. One or more attributes will be associated with each facial feature. Some attributes will relate to the size, orientation or position of the feature, others to its shape qualities.

- 2) for each attribute of each feature, the range of variation for that attribute will be represented, and the value of that facial attribute for a given face is given by an interpolation coefficient. Each facial attribute is thus assigned a modifier (the equivalent of adjectives in natural language), but it will not be limited to discrete categories.
- 3) A set of attribute-value pairs for the entire feature set will constitute a description of the overall face.

While language-like in its extensibility, the design of the facial feature set and its component attributes is nontrivial. As noted earlier, modifiers can be used in combinations that have undesirable consequences. The current scheme is not immune to this problem. The primary issue will be one of orthogonality, which enters into all aspects of the descriptive scheme, from choice of feature set, choice of attributes per feature, and how these attributes are explicitly represented.

#### 4.10 Visualising Attributes

To describe some attribute as a fractional interpolate between two extremes would appear to dodge the question, for how then are the extremes defined, such as the exemplars of extreme leptorrhiny and platyrrhiny? How would a given nose be represented as 37% of the way between those two extremes?

The solution adopted here is to provide sculpted exemplars for the two extreme shapes, and to view a third shape (call it the ‘base shape’) that is an interpolation (e.g., 37%) between those extremes. The same base shape would also be adjusted to simultaneously depict many other independent attributes, resulting in a model of the entire face. It might seem ironic to base a visual description ultimately on sculptures, given that the introduction states:

A ‘descriptive visual representation’ is more than just a model or a replica. Even if a three-dimensional model were to faithfully replicate the facial features of some ethnotype, its characteristics would remain *implicit* — the model would not

serve as a description any more than would a sculpture in marble or other physical medium.

To provide a simple example, consider describing a set of cylinders which are identical but for their height, which can vary from 10 cm to 100 cm. Any specific cylinder could be described by specifying its height either in absolute (`height : 55 cm`) or in relative (`height : 0.5`) terms, where  $0.5 \cdot (100 - 10) + 10 = 55$  cm. Then, adding another degree of freedom, suppose the diameter could vary from 2 to 50 cm, requires another attribute-value pair. The overall description of the cylinder increases with the number of descriptors of course, but our ability to appreciate the cylinder based on this description diminishes.

A visual representation, however, is better able to scale as the number of degrees of freedom increases. Starting again with one degree of freedom, height, this could be visualised by observing a morphable model (one that can assume any height within some range). The shape of the ‘base’ model is then computed as a linear interpolation (or ‘blend’) between the shapes of two hidden copies of that model, where one copy has the least height, and the other, the greatest height. The range of heights is implicit in the difference between the two hidden copies. The observer sees only the resultant cylinder with the interpolated height (in the above case, a 50% interpolation). But this seems to get us no closer to a visual representation than if one were to simply provide a scale model or replica of a cylinder of that specific height. The advantage comes in our visual appreciation of *combinations* of attributes, and in comprehending the influence of each attribute separately upon the resultant shape.

In order to add the diameter attribute, another hidden pair of copies of the model is created, one with the smallest diameter and the other the largest diameter, and now the shape of the visible ‘base’ is computed as the linear-weighted combination of the four hidden copies, given two parameters to specify the fractional height within the height range and the fractional diameter within the diameter range.

Incidentally, the interpolation could be simplified by modelling the base shape as having *both* the least height and the least diameter, then providing only two hidden copies, one with the greatest height (but still the least diameter) and the other with the greatest diameter (but still the least height). This approach is less extensible, however, since one would have to start over to accommodate an additional shape attribute. In this study, it was preferable to start with some representative shape (neither the extreme of its attributes nor the mean of its attributes) and to incorporate additional attributes by hidden representations of their extremes. In the case of the simple cylinder, that would suggest starting with a cylinder of arbitrary height and diameter, then to influence that shape as a linear weighted combination of the extremes of those two attributes.

What makes such a dynamic visual representation actually useful is that we can understand the attributes by *watching them in action*. A simple interface could have two unlabelled sliders, one controlling the height and the other controlling the diameter. These two attributes would be defined purely visually, and their meanings would become clear through interaction, despite the absence of a label for each slider. No invention would be required to implement such a visual representation of shape. The deformation of a shape by linear-weighted combination of copies that present variations on that shape is provided by the well-established technique of blendshape deformers which can distort by translating the vertices that comprise a polygonal representation of the shape (Section 6.3). In the application of this technique to this study, facial geometry will be represented by a polygonal mesh, the shape of which will be deformed through a set of control parameters, each parameter representing the value of a corresponding facial attribute. The model, controlled by these parameters, would not only replicate the facial features of some ethnotype — as might a sculpture — the set of attribute values themselves would constitute an explicit description of facial shape.

#### 4.11 Defining a Face Space

The human face exhibits ethnic differences geographically, as well as individual differences within each ethnicity. By what means are ethnic differences to be described? An-

thropometric studies have measured the distances between facial landmarks for various ethnicities and found that few specific dimensional measurements reliably distinguish between even very different facial types (Section 2.2). Only a few facial features are diagnostic of one or another ethnicity by their presence or absence. For example, the epicanthal fold is present in Asian eyes and generally absent in Europeans and Africans, but most other facial features differ only by degree across ethnicities. Virtually every aspect of a face is subject to variation, both among individuals within a population and across geographically-distant populations. In other words, the dimensionality of the space of possible faces (however it is parameterised) is very great. Attempting to appreciate the differences in such high-dimensional data, one encounters many of the challenges common to data visualisation in general.

Since faces vary simultaneously along many dimensions, a representation that spans the range of ethnicities would constitute a multidimensional ‘face space’, where the specific choice of dimensions for this face space depends on the specific application. Two types of face space can be distinguished: an identity face space (IFS) for representing individual variations within a homogeneous population of faces, and an ethnicity face space (EFS) for representing variations of faces across ethnotypes. Both are similarity spaces (Section 2.4), but otherwise there is little reason to expect they share dimensions or indeed dimensionality. In our application of an EFS, we are not concerned with ethnicity recognition (the analogue of IFS-based face recognition), but rather, ethnicity description. A representation that supports efficient automatic recognition is not necessarily well-adapted to guiding a human observer. To explore facial variation across ethnicities, any proposed representational scheme needs to span the range of ethnic variation and to facilitate visualisation of similarities and distinctions between ethnicities.

The concept of an average face is central to an IFS: it constitutes the origin of a space in which to map individuals as variations on a mean (Section 2.4.2). A core presumption of an IFS is that the sample set of faces is homogeneous and that individuals are normally distributed about the sample mean along each IFS dimension, placing the average face at



the origin of the IFS. Since our EFS is not intended for ethnicity detection (the analog of face detection with an IFS), there is no need for a central-tendency presumption. Moreover, we avoid problems of measuring a global mean on which to map ethnicities as variations. A global mean across all facial ethnotypes would exhibit very large variance, since within-ethnicity variance in facial dimensions has been shown to obscure most across-ethnicity differences, even between very different selected ethnotypes. Finally, in not attempting to measure a global mean, one avoids issues of over-sampling, as some very populous ethnicities would greatly overshadow other important but smaller ethnicities. Fortunately, for our purposes, ethnotypes need not be described as variations on a global sample mean of ethnicities.

In this study, an EFS will be restricted to ethnic variation visualisation, as opposed to automated ethnicity detection. An EFS has two core characteristics: 1) it represents faces with absolute rather than relative EFS dimensions, and 2) the individual dimensions semantically meaningful, allowing the space to be ‘navigable’ by humans. In both regards the EFS is starkly different from an IFS, for which the dimensions correspond to relative differences, and being the eigenvectors determined by principal components analysis (PCA) on some training set, those dimensions are ‘holistic’ measurements derived across the entire face (e.g., ‘eigenfaces’) and not semantically meaningful. In contrast, anthropometric face measurements are both local and individually interpretable. In using such measurements as dimensions of an EFS, there is justifiable concern they would be ad hoc and arbitrary. This could be said of many representational schemes that are built upon human-interpretable attributes. Conventional anthropometric measurements are similarly intuitive (e.g., the width of the nose), and reflect what can be efficiently and reliably measured, rather than any arguably complete, principled set of measurements. Likewise, the sample points for image registration and delineation are based on convenient and conventional image landmarks. Issues of arbitrariness, bias, and mathematical completeness are largely avoided by acquiring dense measurements. Moreover, PCA on those large datasets objectively reveal the principal dimensions of facial variation that are free of subjective interpretation. However, this study is concerned with visualisation of facial

differences for a human observer in terms of specific local facial properties, which, while ad hoc, help in the description of human facial variation.

The specific choice of facial attributes to parameterise is arbitrary and in fact open-ended. Each choice of parameterisation defines a different face space, i.e., a different scheme for representing facial shape. A representation capable of capturing subtle facial variations will necessarily have many parameters — intractably-many to be used as the dimensions of a face space. Therefore a broad goal of this study is to first define a sufficiently fine-grain set of descriptive attributes that adequately capture different ethnotypes, then to subsequently evaluate different schemes to compare ethnotypes by navigating simplified face spaces. That is, by adjusting a parametric model such that its shape differs insignificantly from that of a given subject ethnotype (as measured by the displacement between their two surfaces), the modelling scheme would have served its primary descriptive purpose. Then, given models of various ethnotypes (each a set of attribute-value pairs) their facial differences can be explored in terms of their parametric differences. For example, by alternating from one to the other set of parameter values, the mesh representing the face would shift appearance between the two ethnotypes. Interpolation of all parameters in parallel would result in a smooth continuous blend from one to the other face. Alternative schemes for how to transition between the two sets of attribute values correspond to different strategies for navigating from point to point in face space.

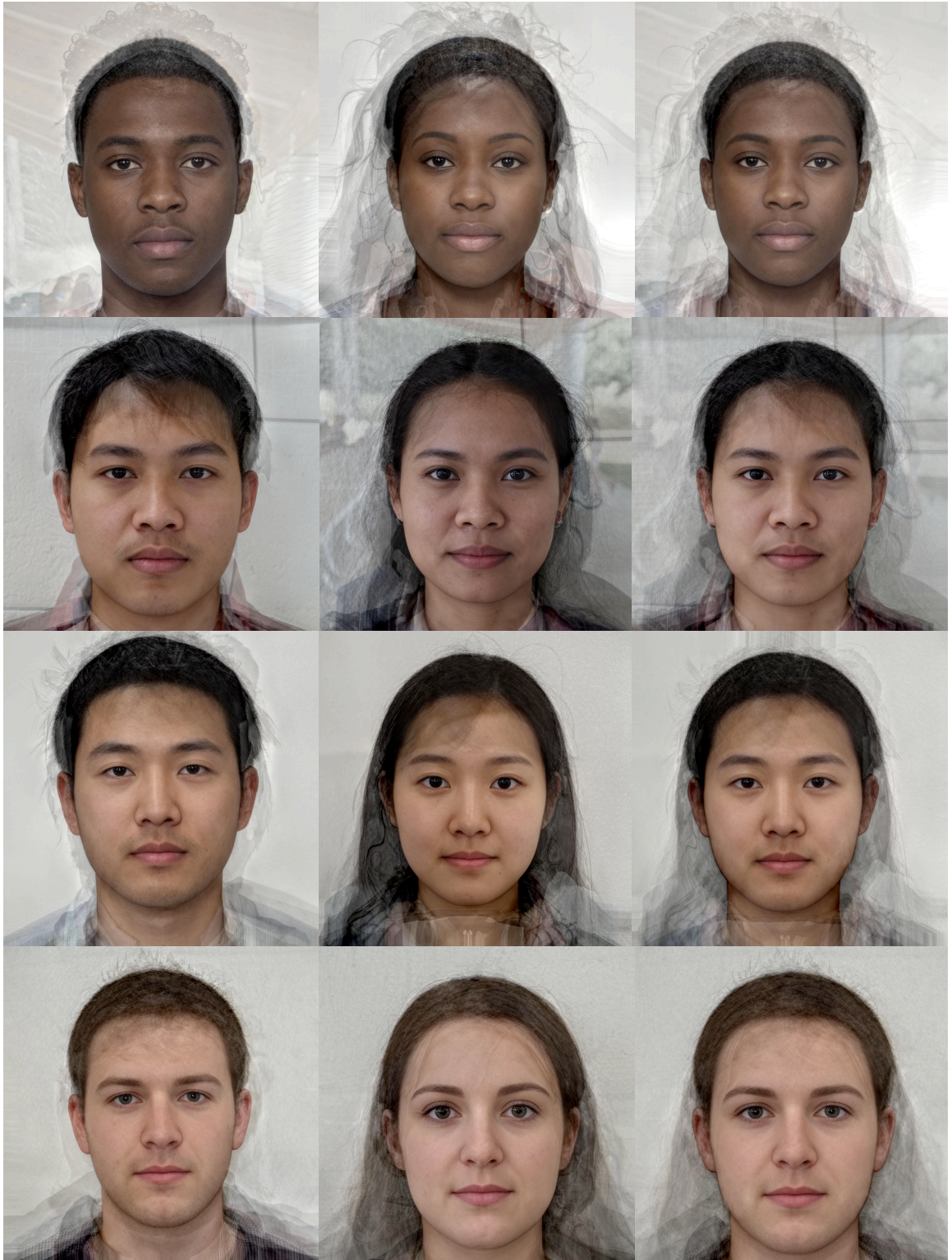
## 5. Facial Features and their Attributes

Gender differences in the human face have been experimentally manipulated by image averaging and blending (Rennels et al., 2008; Scott et al., 2010; Holzleitner et al., 2014). Figure 5.1 shows the averaged faces of four ethnicities (AFR, SEA, EAS, and AFR) for male (left column), female (middle column), and the 50% interpolation between the two sexes. Some facial regions show greater sexual dimorphism than others, for some ethnicities. This study, however, focusses on fundamental ethnic differences without addressing the ethnic aspects of sexual dimorphism. The data and modelling will be specific to the male face, but both sexes are shown for reference in the rest of this section.

In the following, the ethnically-important attributes of each facial region are identified. There are relatively fewer anthropometric landmarks and facial features across the cranium, mid-face, and jaw regions compared to the greater density of features associated with the regions of the eyes, nose, and mouth. The identification of facial attributes therefore began with the eye, nose, and mouth regions, followed by an analysis of how they relate to the overall facial configuration of the cranium, the mid-face, and the jaw. The associated modelling process (Chapter 6) was, unfortunately, not nearly so linear. Many revisions were required in which the underlying polygonal mesh was modified to add or remove vertices in order to adequately capture the extremes of a given attribute or to minimise interactions with other attributes. The search was initially one of identifying ethnically-relevant facial features, but as they were identified, their concrete implementation in terms of basis shapes dominated the process. This chapter describes the resulting catalogue of facial features and their attributes. The modelling of these attributes is then addressed in Chapter 6.

### 5.1 Eye Attributes

As discussed earlier (Section 4.1 and Figure 4.1), the eyes in isolation provide evidence of ethnicity. In Figure 5.2, the averaged images of four ethnicities can be compared (sex differences can also be observed, but our focus is on ethnic differences). While some obvious dimensional differences have been found in anthropometric studies (such as



*Figure 5.1. Averaged faces of four ethnicities (top to bottom: EUR, SEA, EAS, and EUR), for male (left column), female (middle column), and 50% combined sexes (right column). This study focusses on the male face across ethnotypes, recognising that there are also sex differences within each ethnotype. Photographs by L. DeBruine, Face Research Laboratory, University of Glasgow, and the author.*



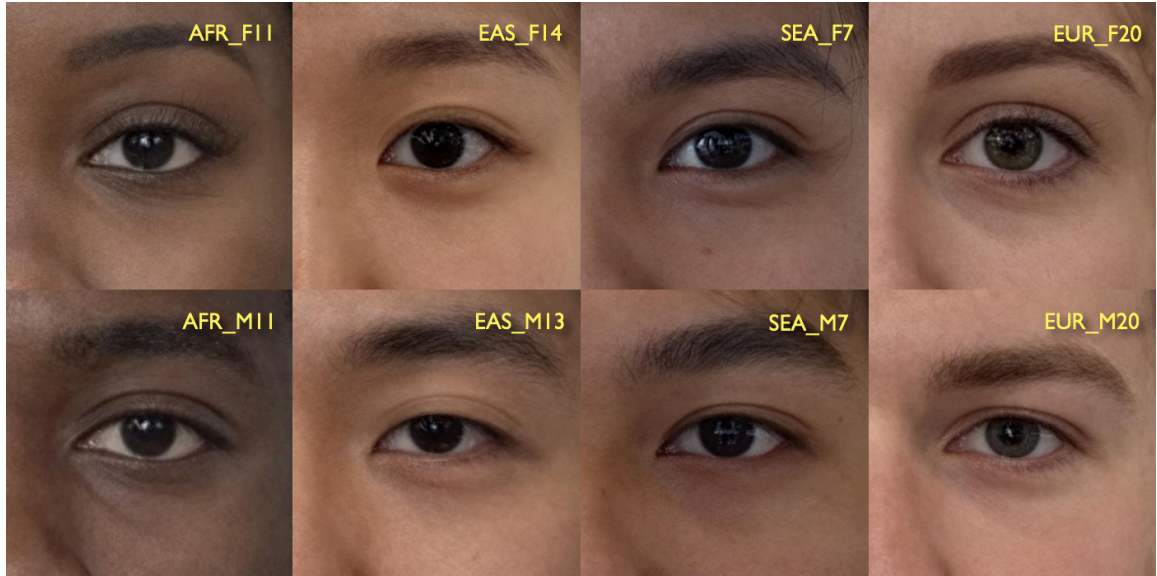


Figure 5.2. Closeups of image-averaged eyes of four ethnicities, for both sexes. *AFR\_F11*, for instance, is the average of 11 female AFR.

palpebral fissure length *en-ex*, Table 2.2), the eyes of these four ethnicities are subtly distinct and reflect shape differences at a smaller scale than are measured anthropometrically. Note that the placement and shape of the eyebrows varies with ethnicity but not considered in this study.

The primary anatomical features of the eye that distinguish the average eye of different ethnicities are listed in Table 5.1 and labeled in Figure 5.3. Figure 5.3 is a closeup of an averaged image of EAS, chosen since the *ECF* is characteristics of many EAS and to a lesser extent some AFR (e.g., L and N in Figure 5.4), but not EUR. The other eye features in Figure 5.3 are present to varying degree across all ethnicities. Individual variation within every ethnotype prevents any simple differentiation of ethnotypes according to the presence or absence of eye features. Compare the averaged EAS male eye in Figure 5.3 with those of the five EAS individuals in Figures 5.4a-e.

They all have an *ECF*, as expected, but they vary considerably in virtually all other features. For example, an *STF* is faintly apparent in A, very pronounced in B, but absent in C-E. Also, A has a deep *SPS*, while that sulcus is slightly shallower in B and replaced

Table 5.1. Features of the eye region, with abbreviations, and their associated attributes. See Figure 5.3 for locations.

Feature	Definition	Attributes
<i>ECF</i>	Epicanthal Fold	<i>ECF_weight</i>
<i>ENC</i>	Endocanthus	<i>ENC_angle</i> , <i>ENC_x</i> , <i>ENC_y</i>
<i>EXC</i>	Exocanthus	<i>EXC_y</i> , <i>EXC_x</i>
<i>IPC</i>	Inferior Palpebral Convexity	<i>IPC_weight</i>
<i>IPD</i>	Interpupillary Distance	<i>IPD_x</i>
<i>IPS</i>	Infrapalpebral Sulcus	<i>IPS_weight</i>
<i>PF</i>	Palpebral Fissure	<i>PF_inclination</i> , <i>PF_width</i>
<i>PMS</i>	Palpebromalar Sulcus	<i>PMS_weight</i>
<i>PNS</i>	Palpebronasal Sulcus	<i>PNS_depth</i>
<i>PTF</i>	Palpebrotemporal Fossa	<i>PTF_depth</i>
<i>SPC</i>	Suprapalpebral Convexity	<i>SPC_weight</i>
<i>SPS</i>	Suprapalpebral Sulcus	<i>SPS_weight</i>
<i>STF</i>	Supratarsal Fold	<i>STF_weight</i>

with an *SPC* in C-E. The *IPS* and *IPC* are both prominent in A while both those features

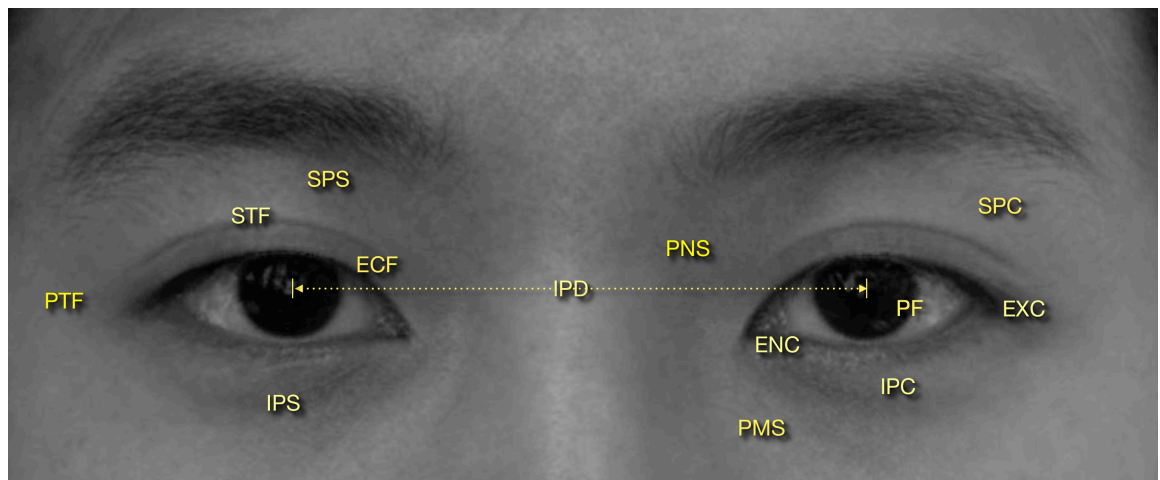


Figure 5.3. Locations of eye features. See Table 5.1 for definitions and associated attributes.

are absent in D. The *IPC* is apparent in B and E, but they are not accompanied with an *IPS*. There is an *STF* apparent in B but not in the others EAS individuals. These differences notwithstanding, A-E all appear obviously EAS.

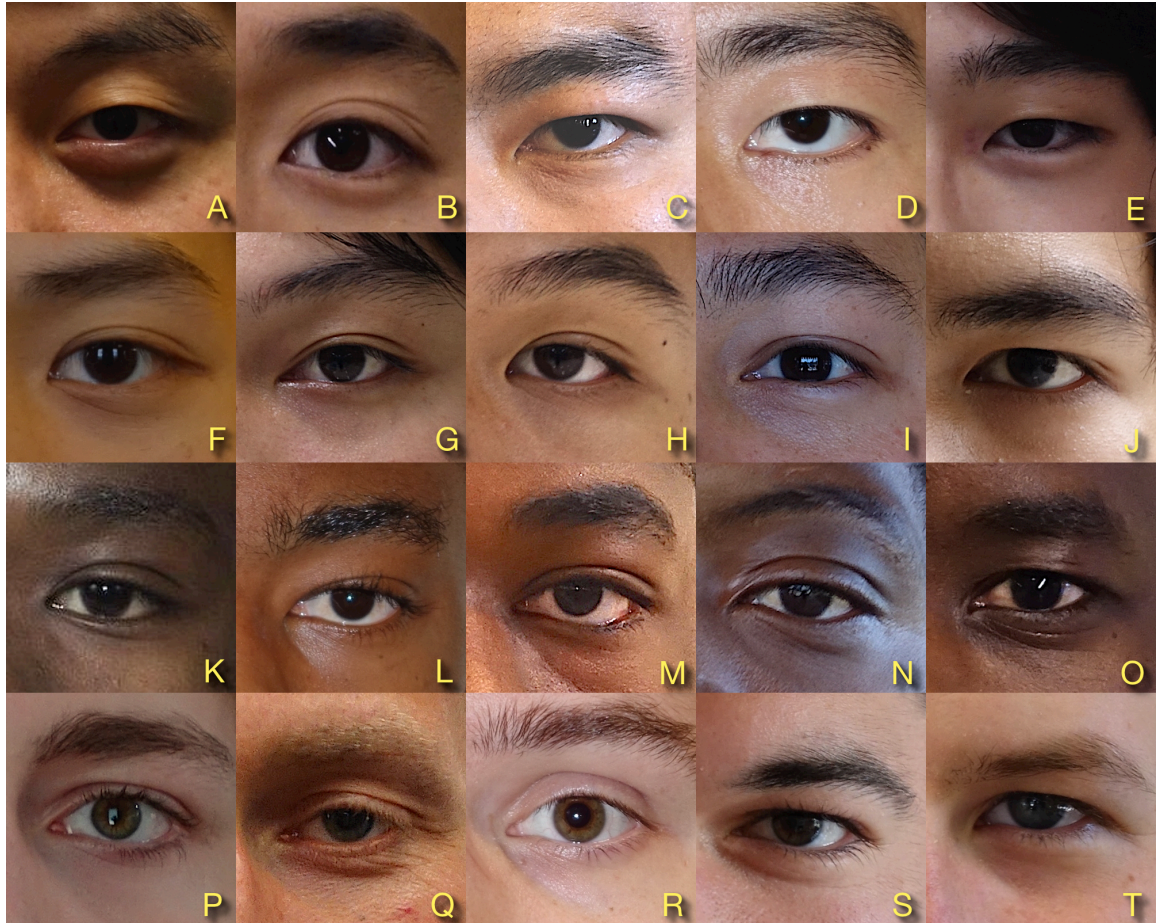


Figure 5.4. Closeups of the eyes of individuals of four ethnicities (top to bottom: EAS, SEA, AFR, and EUR).

Comparing the SEA in the second row of Figure 5.4 with the EAS in the top row, the *ECF* is somewhat less prominent in SEA, along other more subtle differences. The SEA in F-J are ordered with increasing *SPC* (as are the EAS just above). When the *SPC* is less pronounced, an *STF* is often present in SEA. The five AFR K-O show considerable individual differences, as would be expected from such a broad ethnotype. But the eyes of the five EUR individuals P-T are also quite heterogeneous. These eye features can be present to varying degrees — sometimes utterly absent and sometimes quite distinct.

This variability will be represented by associating a normalised value with each attribute. For example, the supratarsal fold will have an attribute *STF\_weight*, where  $0.0 \leq \text{STF\_weight} \leq 1.0$ .



The depth of folds and sulci such as the *STF*, *SPS*, and *IPS*, and the protrusion of the *SPC* and *IPC* are more apparent in side view (Figure 5.5). The supratarsal fold (*STF*) can be seen to occur in eyes with significant *SPC* (e.g., C and L) as well as those with a deep suprapalpebral sulcus (*SPS*), e.g., R and S.

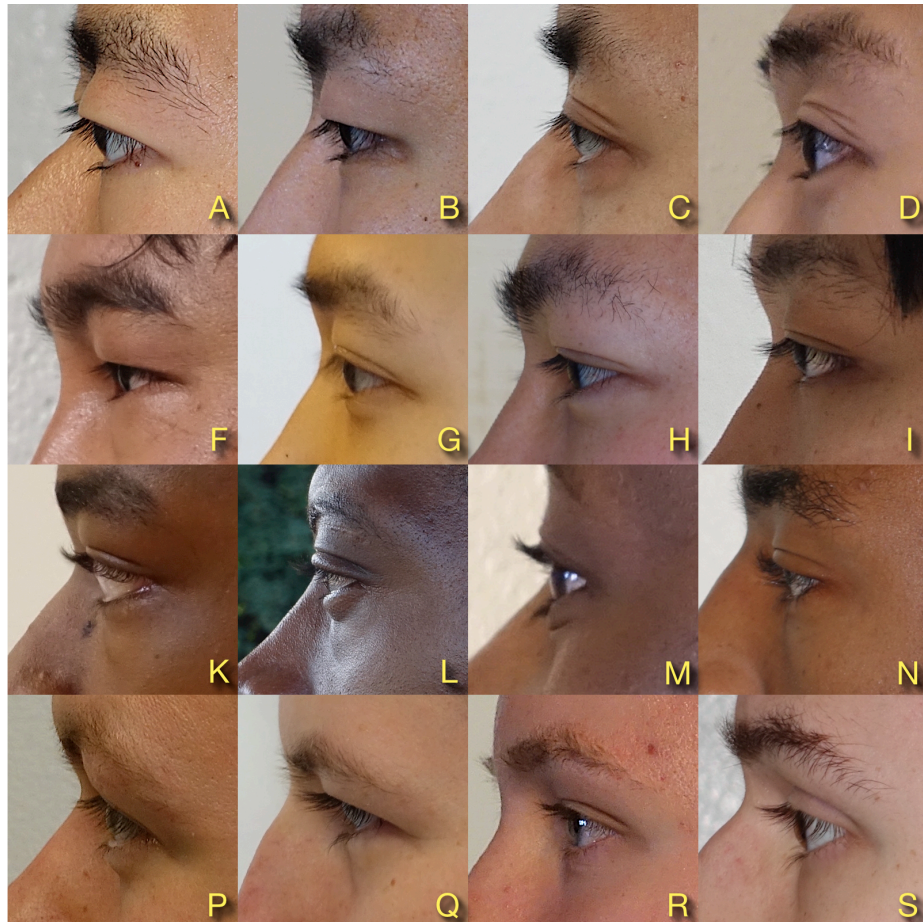


Figure 5.5. Closeups of the eye region of four individuals for each of four ethnicities (top to bottom: EAS, SEA, AFR, and EUR).

The counterparts on the lower eyelid, the inferior palpebral convexity (*IPC*) and the infrapalpebral sulcus (*IPS*), can both be absent (A, F, R) or both can be present (C, G), etc. Again, individual variability precludes any simple, clear-cut rules for discriminating different ethnotypes based on their eye features.



## 5.2 Mouth Attributes

Figure 5.6 shows averaged images of the mouth region for four ethnicities, for both sexes. The salient mouth features are compiled in Table 5.2 and labeled in Figure 5.7.

Nine features are listed for describing ethnic variation in the mouth. This list is not exhaustive; others could be proposed for capturing finer distinctions, especially the particularities of individual mouths, of course. As shown with regard to the eyes, individual variation complicates what might otherwise seem a simple story in Figure 5.6.

In Figure 5.8, four individuals for each of the four ethnotypes are shown, demonstrating that the mouth varies dramatically *within* each ethnotype (often as much as the differences across ethnotypes).



Figure 5.6. Closeups of image-averaged mouths of four ethnicities, for both sexes. *EAS\_M13*, for instance, is the average of 13 male EAS.

Table 5.2. Features of the mouth region, with abbreviations, and their associated attributes. See Figure 5.7 for locations.

Feature	Definition	Attributes
CH	Cheilion	CH_x
CPB	Cupid's Bow	CPB_depth, CPB_width
IL	Inferior Labium	IL_convexity, IL_protrusion, IL_thickness
ILT	Inferior Labial Tuberosity	ILT_fullness
IVB	Inferior Vermillion Border	IVB_curve
PHL	Philtrum	PHL_length, PHL_protrusion, PHL_width
SL	Superior Labium	SL_convexity, SL_protrusion, SL_thickness
SLT	Superior Labial Tuberosity	SLT_fullness
SVB	Superior Vermillion Border	SVB_curve

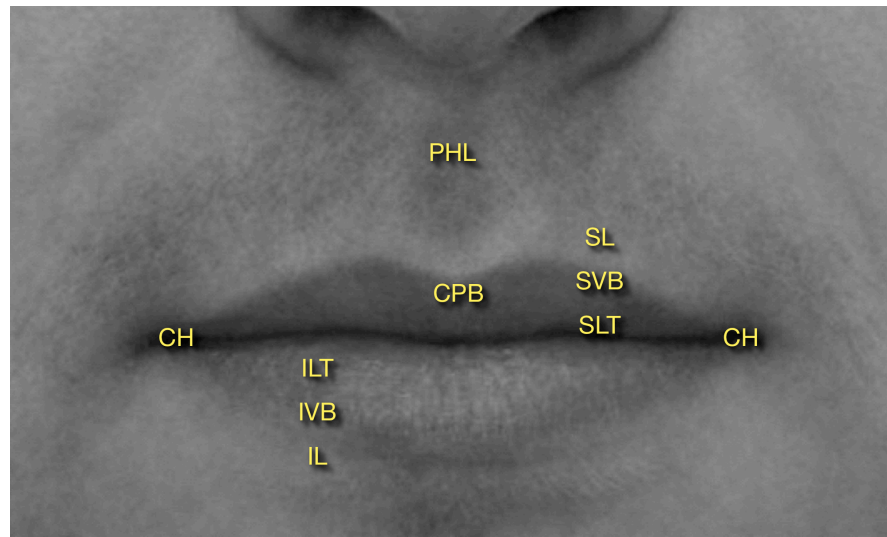


Figure 5.7. Locations of mouth features. See Table 5.2 for definitions and associated attributes.

The shape of the *SVB* (in front view) varies from a simple convex arc (e.g., A) to a re-flex-curve, i.e., convex near the cupid's bow yet concave near the cheilion (G, L, Q and S, especially). This variability in the *SVB* occurs within each ethnotype, to some degree. The AFR mouth tends to have thick and protruding lips, and the *SVB* in AFR is often a raised, sharp crease and not simply a pigmentation boundary as in many EUR. While the ethnicities of individuals A and S are unambiguous, some EAS mouths (G, H, and I) are



*Figure 5.8. Closeups of the mouth region of four individuals for each of four ethnicities (top to bottom: AFR, EAS, SEA, and EUR).*

confusable for EUR mouths (P, Q, and R). Likewise, some SEA mouths are close to AFR in appearance (compare K with C, for example). As was the conclusion regarding eye features, few clearcut rules present themselves for distinguishing ethnotypes based on the features of the mouth.

### 5.3 Nose Attributes

The image-averaged noses of four ethnicities are shown for both sexes in Figure 5.9. Unlike the clear sexual dimorphism in the eye (Figure 5.2) and mouth (Figure 5.6), the differences between the male and female nose is more subtle for each ethnotype.

In front view the traditional distinction (Section 2.2.2) between the platyrrhine AFR, the mesorrhine EAS, and the leptorrhine EUR is apparent, but other than the slightly more bulbous tip of the AFR, the straight ridge of the dorsum in EAS and EUR, the nose does not appear to vary dramatically with ethnotype, at least in front view. The ethnic differ-



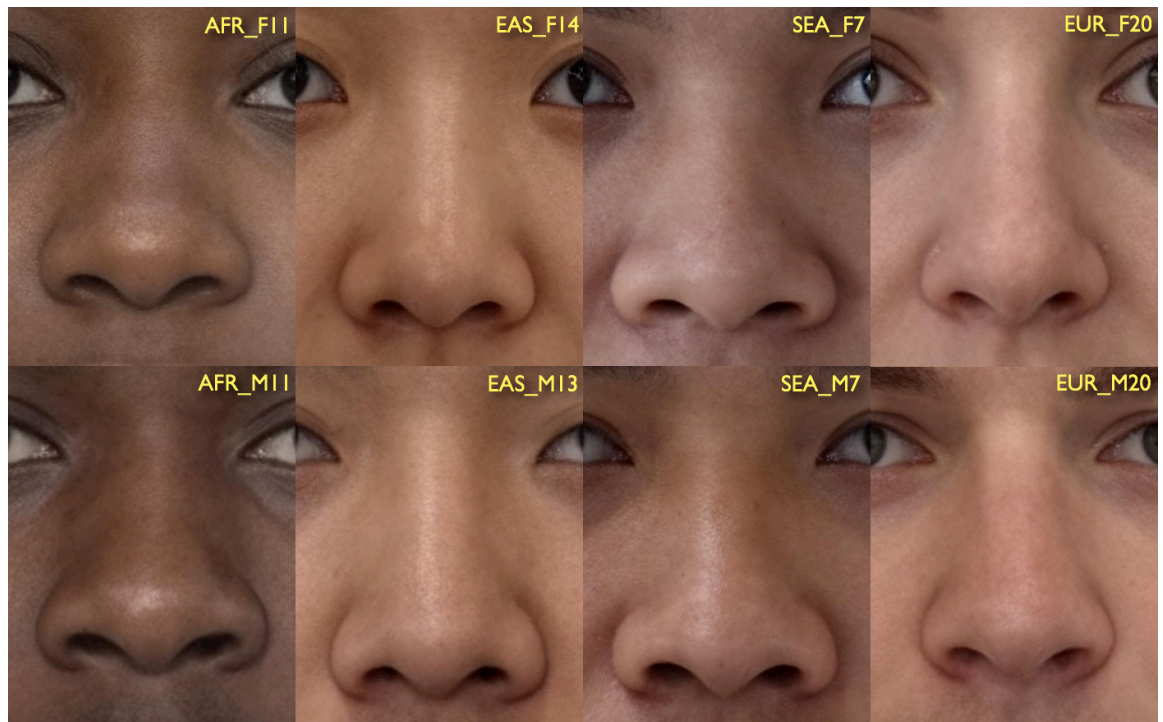


Figure 5.9. The nose region of four ethnicities, for both sexes.

ences are much more apparent in side view. The ethnically salient features of the nose are tabulated in Table 5.3 and labeled in Figure 5.10. The profile of the nose is formed by the relative protrusion of the underlying nasal bone at the radix down to the rhinion, then the transition to the cartilaginous dorsum, tip, columella, and finally into the philtrum, which is regarded as a mouth feature in this study (Swennen et al., 2006). To consider individual variations within an ethnotype, Figure 5.11 shows four individuals for each of the four ethnotypes.

All images are cropped on the right to barely show the cornea of the eye, then scaled to the same height to show the profile from just above the radix to just below the philtrum. Note in Figure 5.11 that the infraglabellar notch (the radix region of the nose just below the glabella) is characteristically very deep in EAS and AFR (Hanihara, 2000) compared with that in EUR. Compare the EAS P and Q and the AFR F and H with the EUR C and D. Next, observe that the profile of the dorsum varies in all ethnicities from concave (showing a supratip break, in A, H, K, L, and R) to straight (D and N, especially), to convex with a prominent rhinion (B, C, G, I, M, and R). Tip inclination varies from steeply

Table 5.3. *Features of the nose region, abbreviations, and associated attributes. See Figure 5.10.*

Feature	Definition	Attributes
ALA	Ala	ALA_contour, ALA_drop, ALA_width
COL	Columella	COL_drop, COL_show, COL_width
DSM	Dorsum	DSM_length, DSM_protrusion, DSM_width
RAD	Radix	RAD_protrusion, RAD_width
SID	Sidewall	SID_slope
TIP	Tip	TIP_drop, TIP_inclination, TIP_pointed, TIP_protrusion, TIP_width

inclined (A and M) to descending (D and I). Likewise, ‘columellar show’, the exposure

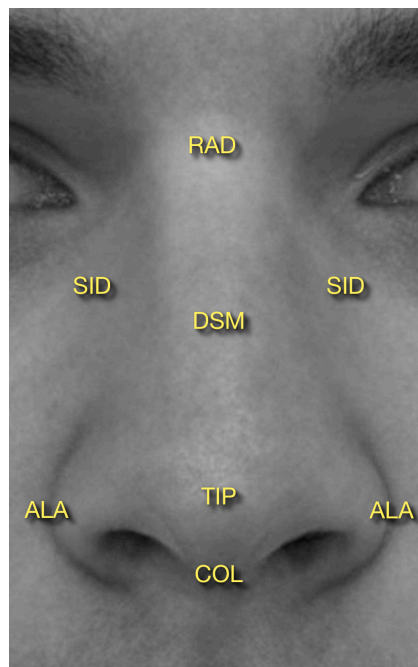


Figure 5.10. *Locations of nose features. See Table 5.3 for definitions and associated attributes.*

of the septum as seen from the side due to elevated alar wings, can be considerable (D, G, and S) or negligible (A, K, and R). Nasal tip protrusion can be far greater in EUR (B and D) than in some AFR (F) and EAS (P). Finally, the relative length of the dorsum (compared to either the protrusion of the tip, or the alar width, or other basis) is especially variable in EUR.



Figure 5.11. Variations in the nose profile in four individuals for each for four ethnicities (top to bottom: EUR, AFR, SEA, and EAS). Not to scale; each profile extends from just above the nasion to just below the philtrum.

#### 5.4 Cranium, Midface, and Jaw Attributes

Next are the cranium, the midface, and the jaw regions. For reference, averaged images of four ethnicities are shown in front view (Figure 5.12) and side view (Figure 5.13). In Figure 5.13 the eyes were aligned to the same height in the figure, and the images were scaled so that all four had the same image scale measured from the exocanthus *ex* (the corner of the eye) down to the cheilion *ch* (the corner of the mouth). This was intended to normalise the face dimensions in order to better reveal different facial proportions for



Figure 5.12. Averaged faces of four ethnicities. Scaled to equal distance from eye to mouth.

the nose, philtrum, mouth, and chin across these four ethnotypes. Note the relatively thinner lips and longer dorsum in EUR compared to those in AFR. The proportion differences (which are quite apparent in Figure 5.13) could be derived from a comparison of the values of the various attributes *DSM\_length*, *PHL\_length*, and so forth, but those proportions are not made explicit (e.g., as indices or ratios).

The origin of the *Ethnicity Modeller*, recall, is placed at the midpoint of a line that just touches the two corneas (see the yellow horizontal line in Figure 5.12). The height (in Y) and depth (in Z) of the cranium CRN can be described relative to this origin. As apparent

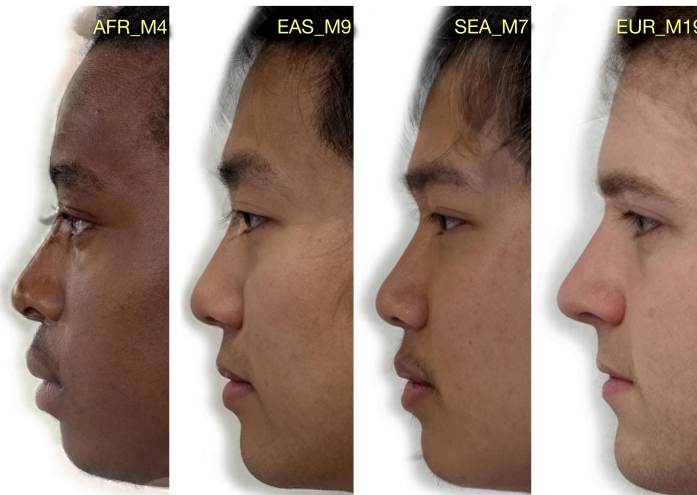


Figure 5.13. Averaged images of four ethnicities in side view, aligned for constant eye height, and scaled to constant distance from eye to mouth.

from the side (Figure 5.13), notice the variations in placement of the supraorbital ridge SOR, the horizontal prominence of the frontal bone under the eyebrow. The height (in Y) and protrusion (in Z) of the SOR varies considerably with ethnicity. Above the SOR, the profile of the forehead FOR also varies with ethnicity in curvature, slope, protrusion, and width — all attributes that share substantial spatial overlap. The FOR, despite its lack of sharply-defined landmarks, does exhibit subtle shape attributes (and, like, other facial regions, the features in this study are not exhaustive: the frontal prominences and the glabella could be added later, for example).

As shown in Figure 5.14, the cranial region is bordered below by the eye region and extends laterally along the temples TMP. The CRN, TMP and FOR all contribute width attributes. The midface region lies adjacent to the nose and just below the eyes, and consists of the maxilla MAX, and cheek CHK, and it extends laterally to include the zygion ZYG and the tragion TRG. The jaw region then extends below the mouth and includes the labiomental sulcus LMS, gnathion GNA, mandible MDB, gonion GON, and neck NCK. These features, labeled in Figure 5.14 contribute the attributes in Table 5.4.

Considering the complete set of features and attributes in Tables 5.1 through 5.4, a total of 77 attributes were identified as salient in describing ethnic differences. The attributes are either dimensional (such as the various widths, lengths, heights, and protrusions) or shape-related (such as the various curvatures, slopes, weights, and convexities). The attributes are intended to be localised and relative to their facial region. Protrusion of the tip of the nose or the gnathion would not correspond to their absolute placement in Z, but instead, would be relative to the nose or jaw, respectively. Likewise the placement of the mouth would be relative to both the nose and the face, and not simply described relative to the model origin. The approach taken here is to create a vocabulary of attributes within facial regions. While local proportions of the nose (e.g. the width-to-height ratio or nasal index  $al-al/n-sn$ ) can be derived from the ratio of the ALA\_width to DSM\_length, that index would have to be computed from the two attributes, rather than being regarded as a nasal attribute. More generally, proportions will not be treated as primitives of this



Table 5.4. Features and attributes of the cranium, midface, and jaw. Refer to Figure 5.14.

Cranium	Definition	Attributes
CRN	Cranium	CRN_depth, CRN_height, CRN_width
FOR	Forehead	FOR_curvature, FOR_protrusion, FOR_slope, FOR_width
SOR	Supraorbital ridge	SOR_height, SOR_protrusion
TMP	Temple	TMP_width

Midface	Definition	Attributes
CHK	Cheek	CHK_protrusion, CHK_width
MAX	Maxilla	MAX_protrusion
TRG	Tragion	TRG_y, TRG_z
ZYG	Zygion	ZYG_protrusion, ZYG_width

Jaw	Definition	Attributes
GNA	Gnathion	GNA_x, GNA_y, GNA_z
GON	Gonion	GON_x, GON_y
LMS	Labiomental Sulcus	LMS_weight
MDB	Mandible	MDB_protrusion, MDB_width
NCK	Neck	NCK_protrusion, NCK_width

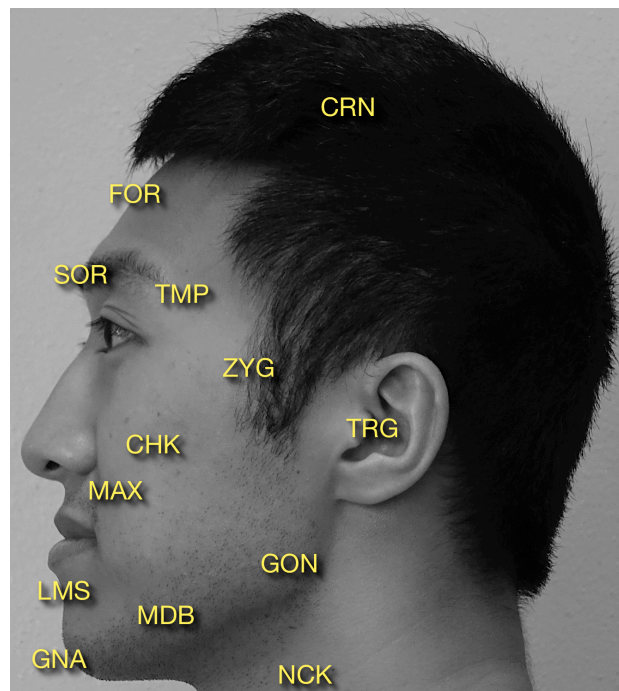


Figure 5.14. Locations of features of the cranium, midface, and jaw. Refer to Table 5.4.

representational scheme.

Having compiled features within six facial regions, and having associated attributes with these features that permit quantitative descriptions, the next step is to implement them in concrete terms, as basis shapes.

## 6. Developing the Ethnicity Modeller

The *Ethnicity Modeller*, or *EM*, that was introduced in Chapter 3 will allow ethnicity visualisation through the deformation of a polygonal mesh. A set of attribute-value pairs (Section 4.2) will serve as a quantitative model of face shape, and the shaped mesh that is deformed according to those attribute values will serve to visualise that face. In the following, therefore, a ‘model’ has two meanings: it will either refer informally to the polygonal mesh as a representation of the surface of human face, or more technically, as a file that stores a complete set of attribute-value pairs.

The core of the modeller is a polygonal mesh called the *base mesh*, that resembles a rather generic human face (Figure 6.1). The shape of the base mesh is simply a default to be subsequently deformed; it is not intended to resemble any ethnicity or average of ethnicities.

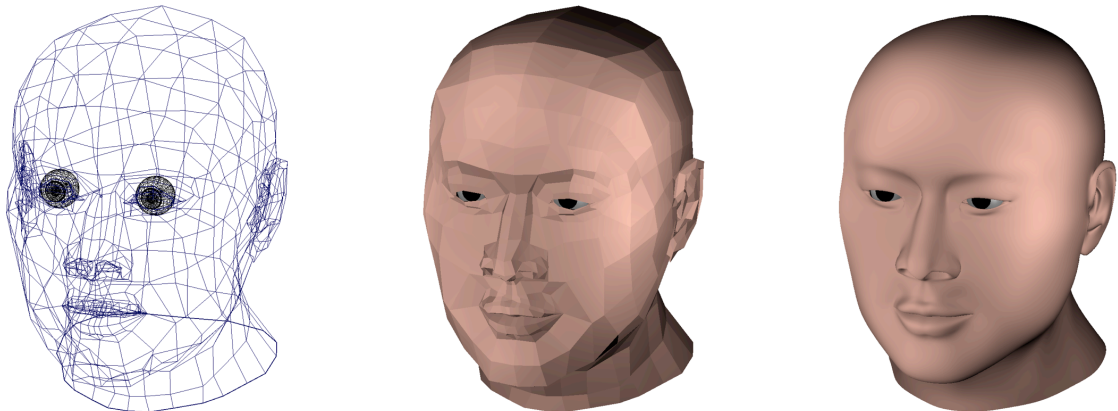


Figure 6.1. The base mesh rendered shown as a wireframe (left), as a polygonal mesh (middle), and as a smooth smooth surface (right).

As discussed (Sections 2.1.3 and 2.4.2), it would not be meaningful to describe faces as variations on some sample mean across faces of all ethnicities (unlike an IFS based in individuals from a homogeneous population wherein the sample mean does serve as the origin). A user interface will control the deformation through a set of parameters, each adjusting the contribution of a corresponding facial attribute towards the resulting facial

shape. The polygonal mesh was designed to support the expected range of variation across each of the attributes that were defined. This process was necessarily iterative, requiring the prototyping of many generations of meshes of increasing complexity as additional attributes were incorporated into the modeller and tested. As various combinations of attribute values were tested, unexpected interactions were revealed between what were intended to be independent attributes, forcing revision of their definition. The following first discusses the process of defining attributes, then the process of creating three-dimensional representations of those attributes, and followed by discussion of matching the model to real data.

### 6.1 Defining Facial Attributes by Extremes

In the case of an IFS, different individuals project to distinct points, the average face is represented by the origin of this space, and the distance from the origin along any axis corresponds to the deviation of an individual from the average along that dimension. For an EFS, however, there is no single norm, or mean, relative to which different ethnicotypes may project (Sections 2.1.3 and 2.4.2). If one expects an EFS to be a similarity space, then similar ethnicotypes (or individuals of similar ethnicotypes) would project to nearby points in EFS, and dissimilar facial types would occupy distant parts of the EFS. There would be no expectation that any particular ethnicity would occupy the origin, however.

In this study, the goal is to represent the dimensions of variation in an ethnicity face space or EFS that is immediately useful for appreciating ethnic differences (not automated ethnicity detection). The EFS dimensions are semantically meaningful and correspond to the attributes of facial features that have salient contributes to ethnic variation. Since this work comes *prior* to any principal components analysis of these dimensions (to reveal which attributes are relatively more important than others), each dimension is regarded equally (none are ‘more principal’ than others). That resulted in an EFS with a large number of dimensions. While that space would be intractably large to ‘manually’ navigate one DOF at a time, this study is not concerned with creating the most computa-

tionally-efficient, reduced-dimensionality EFS. Instead, the goal here is to create a sufficient-complete, semantic, and intuitive face space that reveals ethnic variation.

A dimension in this EFS corresponds to some attribute that is salient to the description of ethnicity. Without attempting to determine what the mean value of that attribute might be across all faces, the attribute is defined by its extremes instead. What constitutes an ‘extreme’ of an attribute depends on the nature of the attribute. Some attributes would be simply dimensional, such as the width of the cranium or the length of the philtrum. The extremes for dimensional attributes can be estimated from the anthropometric literature to find bounds within which different ethnotypes would be expected to lie. Other attributes would correspond to landmark positions, such as the location of the gonion or the zygion, and also expected to fall within bounds that have been determined by anthropometry. Again, for a given ethnotype, *where* the value of a given attribute lies on this axis between those extremes is a matter of discovery after the axes are defined. Finally, novel to this study is the attempt to isolate attributes that are not specifically dimensions or locations, but shapes. These attributes would include the properties of features such as curves, creases, hollow regions (deep sulci and broader fossae), and protuberances or convexities. These features may also vary in degree (e.g., from non-existent to subtle to pronounced).

The next fundamental issue to consider is whether an attribute should be regarded as having two extremes, or just one. One extreme would be applies if an attribute varies from absent to maximum, such as in the case of an epicanthal fold in the eye region, which is utterly missing in some ethnicities, and very pronounced in others, with intermediate degrees present as well. This range would be described by an attribute that varies from 0.0 (absent) to 1.0 (most extreme). These attributes will be termed *unsigned*. In contrast, dimensional, positional, and shape attributes are *always* present, however, and would vary in degree, from -1.0 to 1.0, which will be termed *signed*. Note that once the bounds are defined, the midpoint in that range would indeed define a zero, but that would not be

expected to correspond to a statistical mean for that attribute. It is nonetheless worth examining once the attribute system is designed and implemented.

## 6.2 Facial Regions

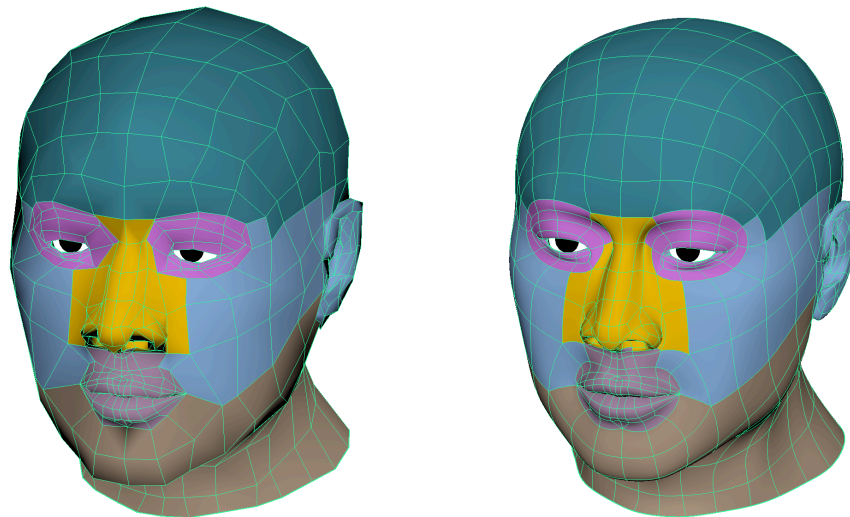
The overall shape of the face is represented as a smooth surface to be shaped locally, feature by feature, according to the specific values of the attributes that it will be given. Following standard practice in digital modelling, a polygonal mesh will be used (Figure 6.1). While faceted, the mesh can be rendered as a subdivision surface according to standard practices to appear smooth (Figure 6.1). The simpler the mesh, the smoother the resultant surface, which reduces the distraction of viewing a faceted, unrealistic surface representation. Yet sufficient topographical detail must be provided in the polygonal mesh to permit the modelling of features (sulci, folds, etc.) that are perhaps prominent in one ethnicity and yet diminished or absent in another. This requires extensive refinement of the model's mesh topology so that these features can appear or disappear according to the ethnicity being depicted. In order that a fold or crease can be made to visually disappear, the mesh has to be able to be made completely smooth, which is very difficult to achieve for fine meshes. There is consequently a trade-off: the mesh must be fine enough (the spacing between vertices sufficiently close) to provide sufficient detail to create the features that are present in some ethnicities, and yet not so much detail that it is difficult to have those features effectively disappear into the surface for other ethnicities. Given that some ethnicities have closely-spaced, sharply-folded creases in the upper eyelid, and the mouth varies subtly in curvature along the vermillion border, this requires careful design of the underlying mesh.

It is fortunate that the human face can be modularised into fairly distinct spatial regions as defined by the underlying cranial and facial bones. The eyes are encapsulated by bone, with the supraorbital ridge of the frontal bones the boundary between the eye region and the forehead, and the zygomatic and nasal bones the boundary between the eye region and the cheeks of the face and the nose. The lateral extent of the face and jaw is defined by the zygomatic bones and the mandible over which layers of soft tissues create

a smooth surface. The protrusion of the nose and mouth is defined by the underlying nasal, premaxilla, and maxilla bones.

While the face can be subdivided into various numbers of regions (the nose, for example, can be further subdivided, as can the relatively smooth forehead), this study will define six non-overlapping regions: the cranium, the eyes, the face (the frontal regions below and lateral to the eyes), the jaw, the mouth, and the nose (Figure 6.2). This subdivision is similar to the regions defined for facial segmentation in other studies (e.g., Steyvers, 1999; Blanz and Vetter, 1999; Zhang and Badler, 2006). Each region has minimal spatial overlap with its adjacent regions but the surface is continuous and smooth at the margins between regions.

With few exceptions (discussed later) the attributes in one region have little effect on those of other regions. Some complications will still occur, e.g., in the area between the nose and the inner corner of the eyes, or in the maxilla and the shape of the mouth. These are inevitable consequences of the fact that these facial features are adjacent and



*Figure 6.2. Six facial regions — the cranium, eyes, nose, mid-face, mouth, and jaw — comprise a useful segmentation as few facial features span their boundaries. The face mesh is rendered as polygons (left) and as a smooth surface (right). Vertices along the boundaries of the regions are shared across regions and will be influenced by attributes in the neighbouring regions.*

share an underlying foundation of bone. Where interactions do occur, the consequences are usually consistent with human facial physiology. For example the total protrusion of the tip of the nose depends on both the protrusion of the nose relative to its base (an attribute of the nose), and the protrusion of the maxilla (an attribute of the face). The two attributes are additive, naturally. While the combination of a very protruding nose and a very retracted maxilla of a flat face may be unlikely, it is not excluded.

For the most part, the various facial attributes will be defined within the boundaries of each region, and not *across* regions. Nonetheless, care is required in placing these region boundaries. The supraorbital ridge *SOR* (the prominence of frontal bone just under the eyebrows), for example, is assigned to the cranium region, but it lies just above the eye region, and therefore variations in the *SOR* will necessarily have some affect on adjacent eye attributes. In retrospect, the segmentation of the face into spatially-adjacent regions was useful primarily in organising the study. The grouping of the attributes by feature within region, is reflected in the user interface design as well as in the structure of the *EM* implementation in Autodesk Maya®. But once defined, the attributes work substantially independently in the *EM* implementation.

### 6.3 Implementing Facial Attributes by Basis Shapes

The modelling of the face follows the conventional practice in computer graphics wherein a smooth surface is represented by a sufficiently-fine tessellation of polygonal (triangular or quadrilateral) facets. The polygonal nature of the surface can be deemphasised by increasing the polygon count, by smooth surface interpolation (e.g., subdivision surface modelling), and by shading.

The shape of a polygonal mesh can be deformed by displacing the locations of their individual component vertices. While keeping the topology of edges connecting the vertices fixed, but shifting their relative positions, the same surface mesh can assume different surface shapes, e.g., changing the proportions and curvature of smooth surface features, and creating folds and creases on what was originally a smooth surface. Variations in



surface shape are then readily achieved using conventional practices of applying so-called ‘deformers’ to a mesh representing a given surface. Multiple deformers can be used in combination, all contributing towards deforming a common polygonal mesh. With care to avoid undesired interactions between deformers, a large set of separate deformers can be created to act as a set of independent facial attributes, each controlled by a corresponding parameter in a user interface (Appendix 1). The ‘base’ polygonal mesh that represents the facial surface must have sufficient complexity replicate a large range of ethnicities. If a specific fold of skin is considered an important feature of the eye, for example, sufficient geometry must be provided in the base mesh to create that fold by shifting of vertices in the vicinity. Through the careful use of deformers it is possible to implement each attribute that is identified as important to this study, to modify the breadth of the bridge of the nose, the curvature of the upper lip, the slope of the forehead, and so forth. The effect of each deformer should be isolated so they act independently, serving as orthogonal facial attributes. Sufficiently many deformers must be provided in order to capture the significant facial on across ethnicities, yet small enough to constitute a descriptive ‘language’ — an encoding of these variations in facial shape.

The magnitude or effect of each deformer is controlled by a normalised parameter associated with that deformer. The degree to which that deformer displaces vertices in the base mesh can be a linear interpolate of the undistorted state and that of the target mesh. A given ethnicity will then be represented by the specific set of parameter values. Through the use of parametric deformation, the model to be used as a descriptive tool, where the facial shape of a given ethnicity is captured by the set of parameter values (or attributes), as well as a visualisation tool.

For this study, the specific set of parameters will be visually-salient features of the human face which readily appreciated and compared. The work is intended to assist humans in visualisation, and thus it would be of little value to create a model that, while successful in capturing different facial shapes, is based on parameters derived from global measurements, as might a set of principal components of shape variation that would be

suitable for a computer vision algorithm. Emphasis will therefore be placed on *local* shape features, since human vision is particularly sensitive to shape differences over small regions (as in the shape of the upper lip, the curvature of a nostril, or the slant of an eyelid). While holistic or global properties (such as derived by principal components analysis, Section 2.4.2) might be valuable for automated face recognition or categorisation, since they are not visually intuitive, they are of lesser value for visualisation of localised facial features.

Ultimately, a representation of the facial shape of a given ethnotype will be replicated by adjusting the values for the deformer parameters of the model. But since individual variation places spatial uncertainty on what might be regarded the ‘mean’ face shape for any ethnicity, and statistical limits on how precisely the surface shape of the parametric model can match that of a given ethnotype.

### 6.3.1 *Meshes and Deformers*

The fundamental technique to change the shape of a polygonal mesh is through a *deformer*, as used in digital character animation to create a large space of possible facial expressions by displacing vertices in a mesh (Bergeron and Lachapelle, 1985; Parke, 1972; Parent, 2012). Digital animation tools such as Autodesk Maya® provide a sophisticated suite of deformers. Those deformers potentially relevant to facial animation include blend, wire, wrap, shrinkwrap, lattice, cluster, point-on-curve, and muscle rigging (Waters, 1987; Bibliowicz, 2005; Ersotelos and Dong, 2008; Orvalho et al., 2012). Of the alternative techniques available, it initially appeared that wire and cluster deformers would be well suited for implementing shape attributes while lattice and other more global deformers would serve to implement dimensional and positional attributes. While they do perform these roles in reshaping a digital face, they do not themselves act as representational primitives. One does not visualise the wire, lattice and related deformers themselves — only their effect on a mesh when they are applied. Also, such deformers are difficult to localise precisely, by their intention. A wire deformer, for instance, will influence the position of mesh vertices along its path with decreasing affect on adjacent

vertices according to their distance from the wire. This results in a smooth deformation, such as a raised eyebrow, or a lip curling into a smile, that smoothly blends into the surrounding mesh. Experimental prototypes showed that these deformers were ill-suited for defining very closely-spaced features, such as the parallel creases and folds around the eye. The deformer that became clearly best for this study, however, is the *blendshape* deformer (see overview by Parent, 2012).

A blendshape deformer associates a base mesh with a set of basis shapes or ‘targets’. The targets are copies of the base mesh with the same topology of vertices and ordering of edges between them but with the vertices of the targets displaced to represent some specific shape (often a facial expression, such as a raised eyebrow or a smile). The term ‘target’ indicates that this basis shape is the target towards which interpolation would converge as the interpolation value approaches 1.0.

In simplest form, consider a base mesh  $b$  and a target  $t$ , which have identical topologies. All vertices in the base will be displaced by a linear combination of their original position in the base and their corresponding position in the target. Suppose the  $i^{th}$  vertex in the base mesh  $b$  has coordinates  $(x_b, y_b, z_b)$  and the corresponding  $i^{th}$  vertex in the target mesh  $t$  has coordinates  $(x_t, y_t, z_t)$ . Under the influence of a blendshape deformer, for blend weight  $\alpha$  (where  $0.0 \leq \alpha \leq 1.0$ ), the coordinates for that vertex in the base will approach that of the target as  $\alpha$  approaches 1.0. The interpolated position  $(x, y, z)$  is given by:

$$\begin{aligned} x &= x_b + \alpha (x_t - x_b) \\ y &= y_b + \alpha (y_t - y_b) \\ z &= z_b + \alpha (z_t - z_b) \end{aligned}$$

This generalises to allow a base mesh to be deformed simultaneously by summation of multiple targets, in fact, each with a separate blend weight. For example, the components of facial expressions (raised left eyebrow, etc.) can be individually modelled by separate target meshes that are then blended in various combination to form broad range of global

facial expressions. In digital animation, facial expressions need to be modulated and nuanced to appear natural during the continuous transitions from one expression into another, i.e., as the weights assigned to various targets shift in continuously-varying combination. Thus considerable effort is placed on controlling for ‘blendshape interference’, i.e., unwanted shape deformations due to the linear summation of blend target weights (Lewis et al., 2005). Those artefacts are smoothed out or removed by the addition of further blendshapes — ‘corrective blendshapes’ — such as might repair the disturbance to an adjacent facial region caused by the adjustment of blendshape weights in a given region, or problems of two blendshapes acting on the same region. Many corrective blendshapes are required to control for blendshape interference. For example, in the animation of a digital character that was controlled by 64 parameters (blendshape sliders), a total of 946 blendshapes were used, most of which were corrective (Raitt, 2004). The intention is to allow an animator to adjust a facial expression through a manageably-small set of independent (orthogonal) parameters, but to meet this goal, many additional blendshapes, hidden from the animator, are added to make the original set of controls act independently.

In using blendshapes as the basis for implementing facial attributes that are effectively orthogonal (so that these attributes may combine arbitrarily), it is not an option to use corrective blend shapes to cancel undesired interactions across attributes. There needs to be a direct one-to-one correspondence between each basis shape and the attribute it represents.

Each signed facial attribute (i.e., one that ranges between two extremes, from -1.0 to 1.0) would consist of a pair of basis shapes that represent those two extremes, and each unsigned facial attribute (that ranges from absent to extreme, or 0.0 to 1.0) would require one basis shape to represent that extreme. The extreme values of each facial attribute directly in terms of a corresponding basis shape. Combinations of facial attribute values will be represented by linear-weighted summation of these basis shapes. Each extreme of each facial attribute will be represented in the abstract by a basis shape and imple-

mented concretely by a blendshape target. The face of any ethnotype is then visualised by subjecting the base mesh to the combined deformation of a large number of targets.

For example, the alar contour (*ALA\_contour*) would be implemented by deforming a copy of the base mesh in that portion of the nose, the rest of the mesh remaining untouched. This is diagrammed in Figure 6.3, where three attributes (*ALA\_contour*, *PHL\_length*, and *ZYG\_width*) are shown, out of the total of 77 attributes.

To implement the attribute *ALA\_contour*, two basis shapes (*ALA\_contour\_min* and *ALA\_contour\_max*) represent the extremes of the alar contour. A blendshape node (*ALA\_contour\_blend*) will deform *ALA\_contour\_min* to become a specified fractional blend (such as 0.45) between itself and that of *ALA\_contour\_max*. That amount of alar contour is then added to the base mesh (*Base*) through a second blendshape node, *combined\_blend*. This blending process occurs for all attributes in parallel.

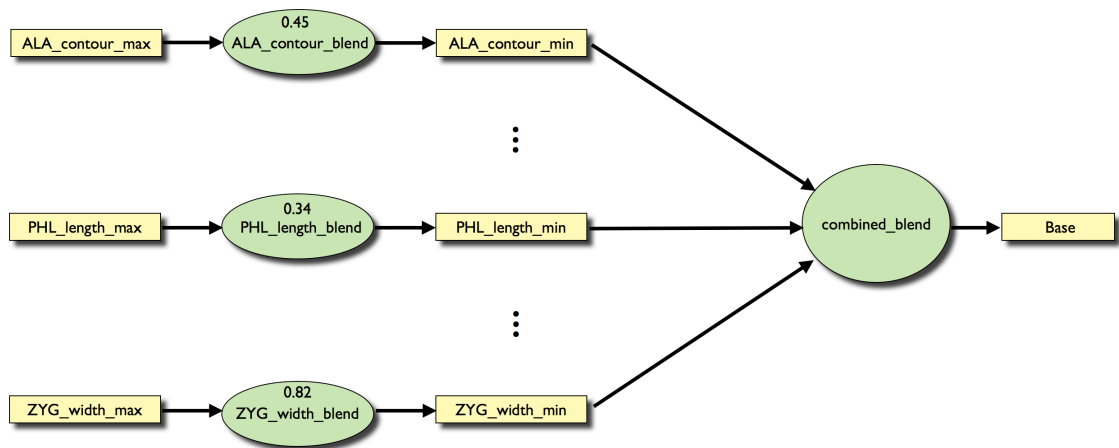


Figure 6.3. The yellow rectangles represent copies of the base (with identical mesh topology but an attribute-specific shape). The green ellipses on the left represent blendshape nodes, each computing an interpolation between the maximum (on the left of the blendshape node) and the minimum (on the right) of that attribute. Next, these deformed meshes, each representing the contribution of an attribute (e.g., a philtrum length of 34%), deform the base mesh in parallel.

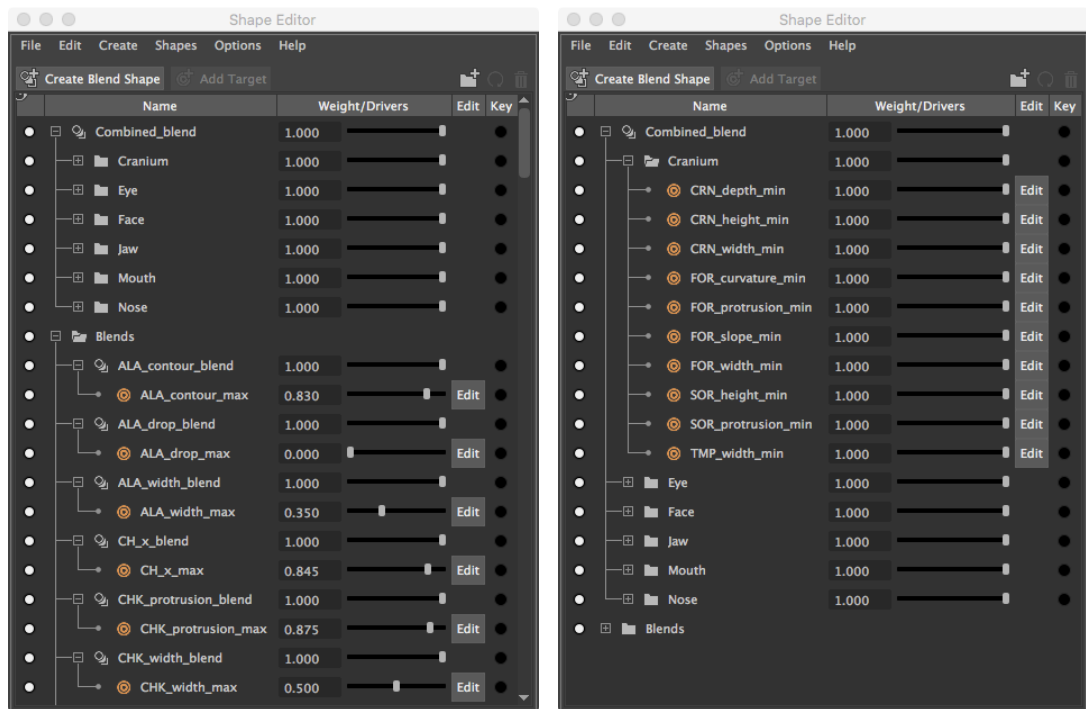


Figure 6.4. Screenshots of the Shape Editor. On the left the *Combined\_blend* shape node is shown as organised into six groups (*Cranium*, *Eye*, etc.). On the right, the *Cranium* group is expanded to show its component targets (*CRN\_depth\_min*, *CRN\_height\_min*, etc.). Below *Combined\_blend* is the group *Blends* (shown collapsed on the right, and expanded on the left) which organises all of the blendshape nodes (*ALA\_contour\_blend*, *ALA\_drop\_blend*, etc.).

Figure 6.4 shows the hierarchical organisation of the blendshape nodes in two screenshots of the Maya Shape Editor. On the left, the targets for the blendshape node *Combined\_blend* are organised into six groups (*Cranium*, *Eye*, *Face*, *Jaw*, *Mouth*, and *Nose*) for convenience. On the right, the *Cranium* group is expanded to show its targets (*CRN\_depth\_min*, etc.).

The Maya Shape Editor screenshots shown in Figure 6.4 illustrate the organisation of blendshape nodes in the *EM*. That editor is not used as the user interface for adjusting the model. Note on the left that the actual blendshape target weights can be seen. For

example, a value of 0.83 for the attribute *ALA\_contour* is represented by a 0.83 as the target weight for the blendshape node *ALA\_contour\_blend*.

### 6.3.2 Attribute Independence and Combinations

In principle, each facial attribute should be able to assume any value within its range with (minimal or no) affect on other attributes. This will be termed ‘quasi-orthogonality’, as it shares some of the properties of ‘orthogonality’ (being ‘uncorrelated’, ‘independent’, and ‘non-overlapping’) as used to describe perpendicular vectors, basis functions, the ‘eigen-faces’ and ‘eigenmeshes’ in principal components analysis and morphometrics, and so forth (Elewa, 2010).

The three-dimensional basis shapes that will represent facial attributes are not strictly orthogonal, since their combined effect will often be spatially overlapping and will contribute to produce common displacements. But many facial attributes can be identified that are substantially independent, even though they may spatially overlap, especially if they are two attributes of one facial feature (such as *DSM\_width* and *DSM\_protrusion*), or attributes of adjacent or partly overlapping features (such as *DSM\_protrusion* and *RAD\_protrusion*).

Orthogonality between two attributes is assured when the two do not overlap spatially, i.e., they apply to different facial features and, in their implementation, do not share mesh vertices. Orthogonality is also straightforward when the two attributes, despite being associated with a common facial feature, are mutually perpendicular, such as the length versus width of the dorsum. Although the two nose attributes share vertices, their implementations displace those vertices in orthogonal (Y versus X) directions, and thus the attributes are effectively orthogonal. But orthogonality is not always straightforward. Many attributes are both spatially overlapping (and therefore share vertices in the mesh implementation) and their displacements are not orthogonal (they produce displacements in the same direction). The two attributes in question might either correspond to a common feature or to two adjacent or overlapping features. For example, one attribute might

correspond to a dimension (e.g., forehead protrusion) and the other a shape attribute (e.g., forehead curvature). Changing one would likely have an effect on the other. These cases are unavoidable since both are salient attributes, and each can be individually present to varying degrees. That is, the feature necessarily has some degree of protrusion, and some shape. These cases cannot be strictly orthogonal in principle, and yet there is utility to allowing their combination. While the two attributes are mutually-inconsistent in some extreme combinations, over most combinations their additivity is plausible in terms of the resultant mesh deformations.

### 6.3.3 The Model Origin

The face model is composed of a polygonal mesh that is placed relative to a modelled pair of eyeballs. The world coordinates such that X is positive to the model's left, and Y is upward, and Z is positive forward. Eyeball dimensions do not vary significantly with ethnicity (Bekerman et al., 2014), hence the eyes were chosen to provide a fixed frame of reference for the face model, since the mesh is deformable and no point on the surface of the face remaining fixed. Specifically, the origin of the model is the midpoint between the anteriormost point on the transparent cornea of each eye, and the two corneas are at  $(\pm \text{IPD}_x/2.0, 0, 0)$ , i.e., spaced equally on either side of the origin with a total separation of  $\text{IPD}_x$ . The Frankfort plane (Swennen et al., 2006) of the model is parallel to the XZ plane. The polygonal mesh representing the face geometry is then sculpted to fit around the two modelled eyes, and remains fixed in space relative to the origin. All mesh deformations will involve only displacing mesh vertices, effectively morphing facial regions in space relative to a fixed model origin. That is, the mesh does not move — only its vertices — during the process of adjusting its shape. The modelled corneas are used as reference points with which to align three-dimensional data meshes (Section 6.5.3).

## 6.4 Basis Shapes for Facial Attributes

A total of 43 facial features were identified in Chapter 5, comprising a total of 77 attributes (10 cranium, 17 eye, 10 jaw attributes, 16 mouth, 17 nose, and 7 midface). This set of attributes is neither mathematically necessary nor sufficient; it is just practically use-

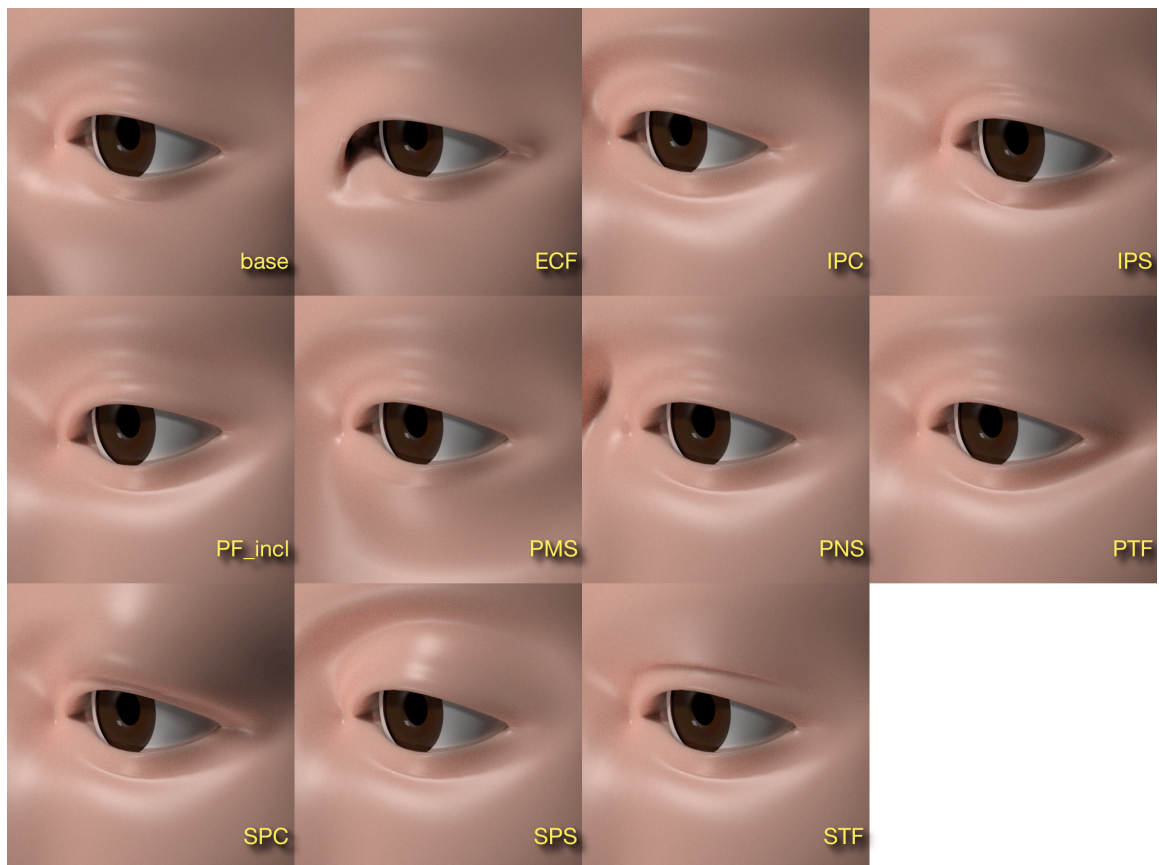


ful, as will be demonstrated. Additional features can be identified and their attributes added as necessary.

#### 6.4.1 Eye Attribute Basis Shapes

A total of 17 eye attributes were modelled by basis shapes. Ten were unsigned (Figure 6.5) and seven were signed (Figures 6.6 and 6.7). In the following, the attributes are introduced individually, and then considered in combination. In Figure 6.5, the base shape of the eye is shown in the upper left, along with the 10 basis shapes representing the extremes of the attributes ECF, IPC, and so forth.

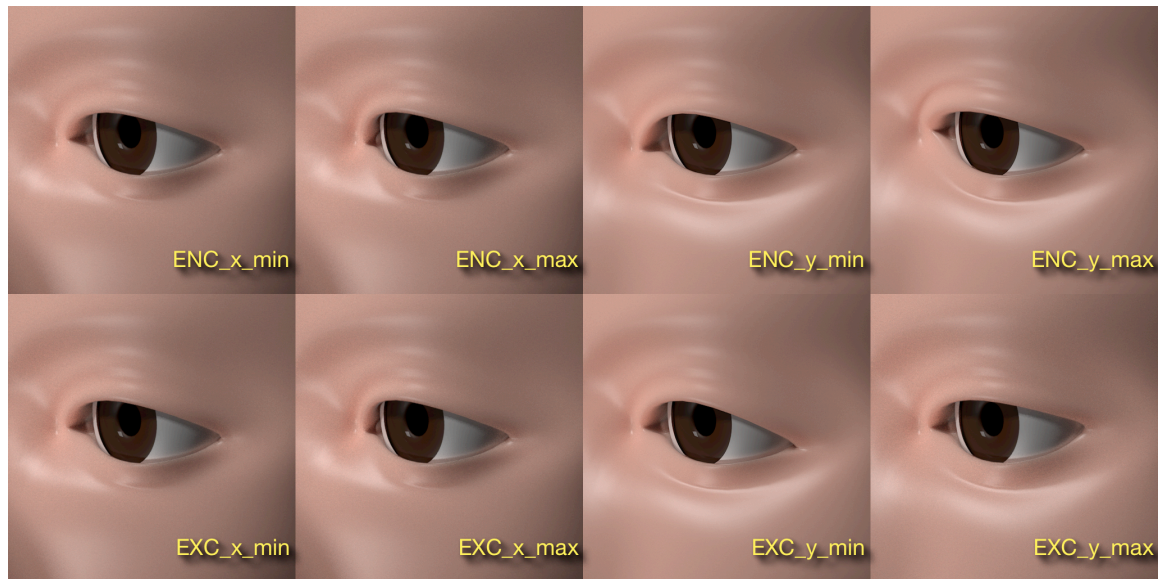
The eye region has a dense arrangement of features. The upper eyelid is particularly challenging to model with a polygonal mesh because the closely-spaced, parallel, and



*Figure 6.5. The base mesh (upper left) plus 10 unsigned attributes of the eye region. Each basis shape represents the extreme of the corresponding attribute (i.e., the maximum deformation relative to the base). Refer to Figure 5.3.*

thin folds of the tarsus (the *ECF* and *STF*) need to vary in depth in the manner of very thin skin tissue, while the adjacent sulci, ridges, and convexities (the *SPS*, *STC*, *SPC*, *PNS*, and *PTF*) need to reflect varying amounts of underlying adipose tissue. A slight and very thin *STF* may overlay the *SPC* (L, Figure 5.4). Significant adipose tissue increases the *SPC* and diminishes the *STF* and *SPS* (C, Figure 5.4), while with very little adipose tissue the *SPC* is replaced by a deep *SPS*, with an *STF* (Q, Figure 5.4) or without (A and R, Figure 5.4). These and many other attribute combinations need to be visualised by different combinations of one set of basis shapes. The signed attributes represent variation in positions, lengths, protrusions, and so forth. These are illustrated in Figures 6.6 and 6.7.

Figure 6.6 shows four attributes, each with the basis shapes that represent their minimum and maximum values. Figure 6.7 shows the other three signed eye attributes, the angle of



*Figure 6.6. The basis shapes for four signed attributes governing the position of the endocanthus and the exocanthus (ENC\_x, ENC\_y, EXC\_x, and EXC\_y). Each attribute requires a pair of basis shapes to permit interpolation within the normalised range -1.0 to 1.0. A value of 0.3 for the ENC\_x, for example, would be represented by a 0.3 interpolation between the basis shapes ENC\_x\_min and ENC\_x\_max. Refer to Table 5.1 and Figure 5.3. See Figure 6.6 for the remaining three signed eye attributes.*

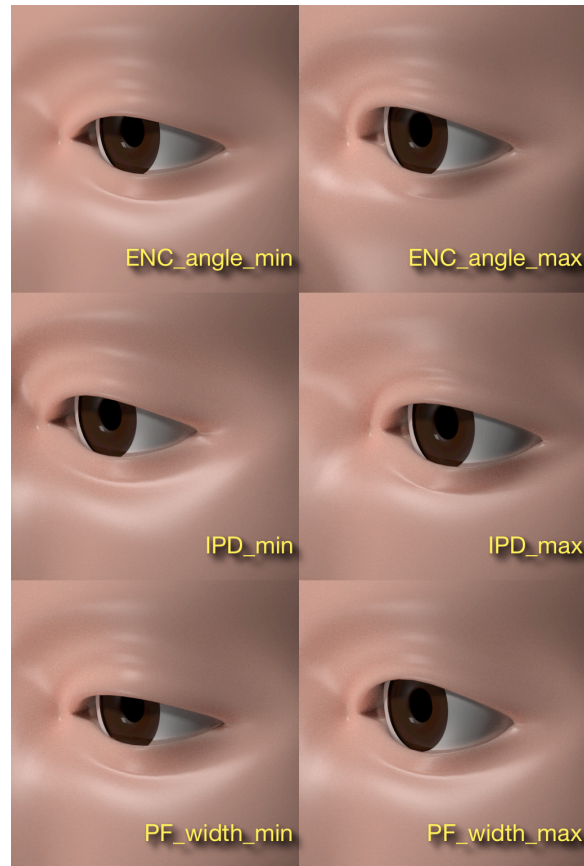


Figure 6.7. The remaining signed eye attributes, each represented by a pair of basis shapes within which the attribute is interpolated. Refer to Table 5.1 and Figure 5.3.

the upper eyelid at the endocanthus *ENC\_angle*, the interpupillary distance *IPD<sub>x</sub>*, and the palpebral fissure width *PF\_width*. Note in particular that the *IPD* shifts the entire eye region laterally within the two extremes depicted. For very narrow-spaced eyes the endocanthus approaches the radix of the nose. The model of the eyeball is displaced automatically in *X* as the *IPD<sub>x</sub>* attribute is adjusted from -1.0 to 1.0. The seven signed and 10 unsigned attributes are able to be combined arbitrarily, as demonstrated in Figure 6.8, and also approximately independent of those of the neighbouring regions (nose, etc.).

These 17 eye attributes can capture a broad range of subtle ethnic variations in the human eye, through various combinations of their values. They have substantial independence in representing folds, ridges, sulci, and fossae which, despite their close proximity and partial overlap, summate in a manner consistent with the observed topography of the skin and flesh surrounding the eye.



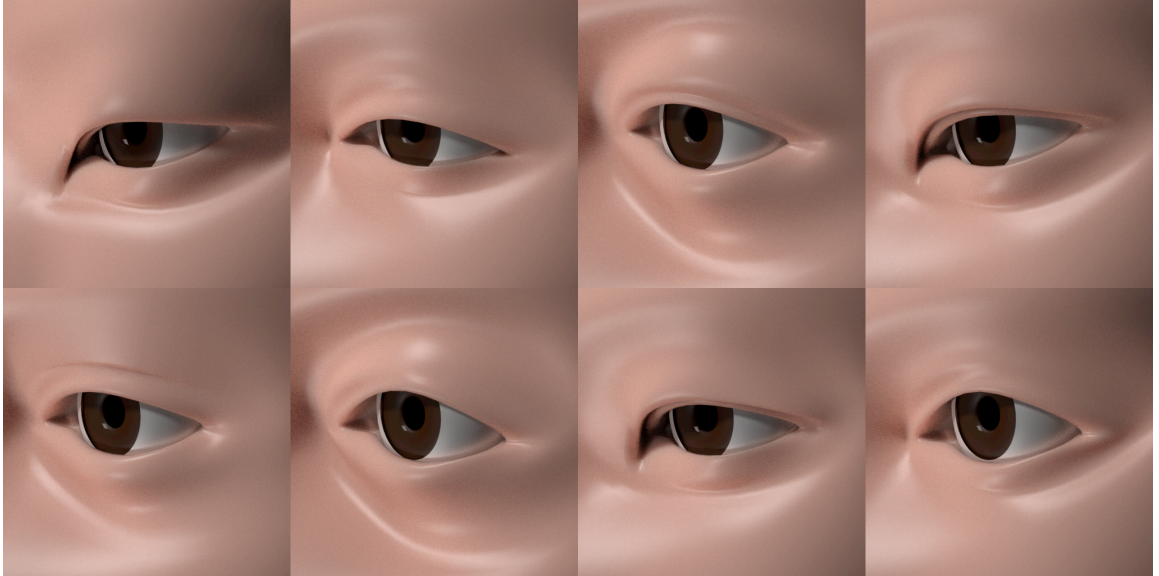


Figure 6.8. Models showing a variety of eye types created by different combinations of the 17 attributes of the eye region.

While this set of attributes captures much of the observed variation, it is not comprehensive. For example, the line-like feature of the upper eyelid, the supratarsal fold *STF*, has only one attribute in this model, namely the *depth* of that fold, but the *STF* also can vary in position (in Y) relative to the margin of the upper eyelid (i.e., ‘tarsal show’) and length. For purposes of this study, however, the current set seems adequate. More attributes would be required in order to extend this approach to modelling individuals, which is beyond the scope of this study.

#### 6.4.2 Mouth Attribute Basis Shapes

Next, the basis shapes representing the attributes of the mouth are discussed. The mouth region has nine features, with a total of 16 attributes. Three of the attributes are unsigned (*IL\_convexity*, *PHL\_protrusion*, and *SL\_convexity*), and shown in Figure 6.9. An additional 13 signed attributes are shown in Figures 6.10 and 6.11.

The features of the mouth have considerably more spatial overlap than those of the eyes. The area of the upper lip alone has five features (CH, CPB, SL, SLT, and SVB) and is

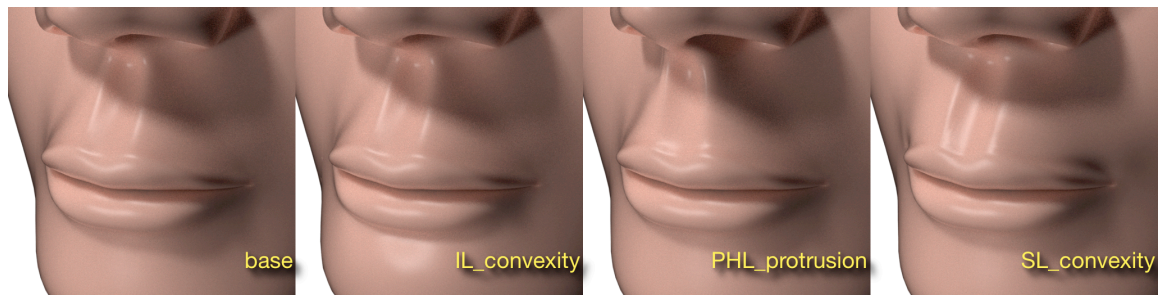


Figure 6.9. The base plus the three unsigned mouth attributes: IL\_convexity and SL\_convexity describe the subtly thicker tissue surrounding the lower and upper lips respectively. PHL\_protrusion is the increased prominence of the philtrum. Refer to Table 5.2 and Figure 5.7.

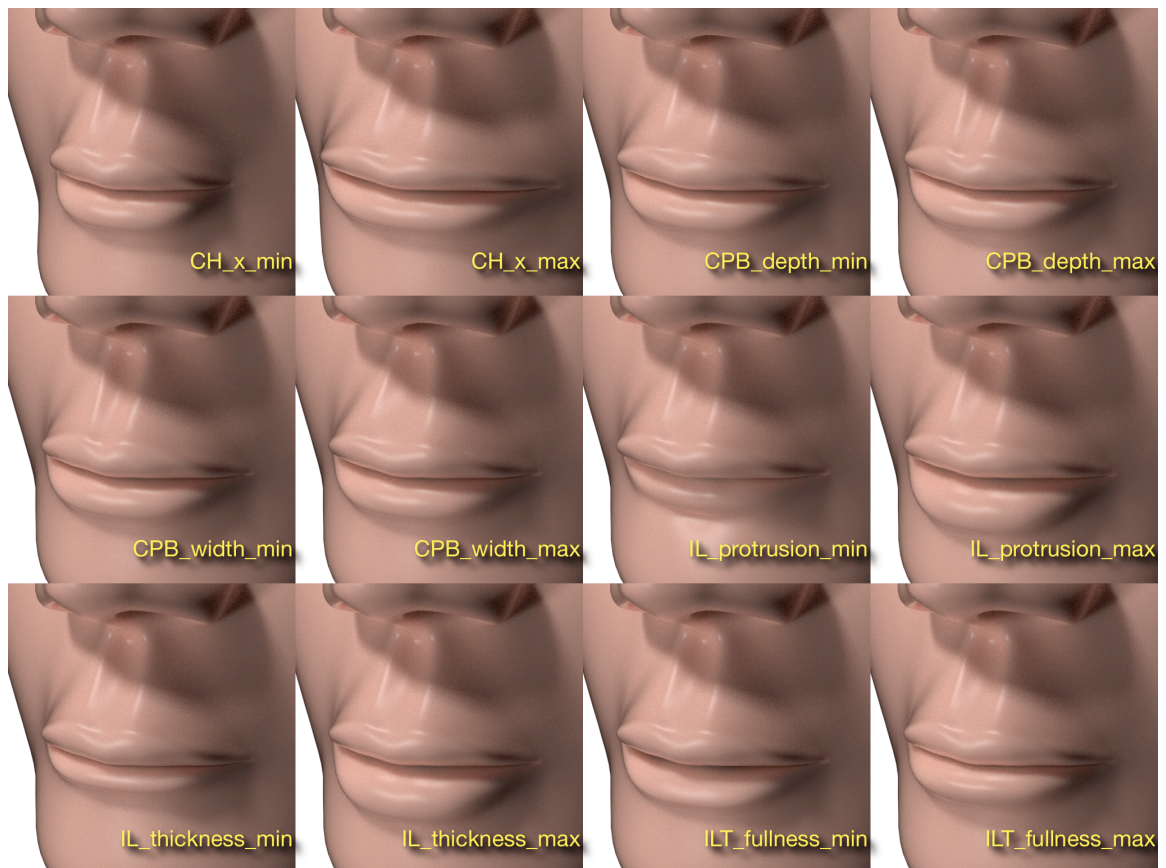


Figure 6.10. Six of the 12 signed mouth attributes, each represented by a corresponding pair of basis shapes. The mouth features are the cheilion CH, the cupid's bow CPB, the inferior labium IL, and the inferior labial tuberosity ILT (refer to Table 5.2 and Figure 5.7).





Figure 6.11. Basis shapes for the other four mouth features: the inferior vermillion border IVB, the philtrum PHL, the superior labium SL, the superior labial tuberosity SLT, and the superior vermillion border SVB. Refer to Table 5.2 and Figure 5.7.

bordered above by the PHL; all overlap considerably (i.e., share multiple vertices), and not simply along their borders. The complexity seems inevitable, however. The superior labium has a vermillion border *SVB* with a very distinct curvature starting at the cupid's bow and ending at the cheilion, the corner of the mouth (Figure 5.7).

The *SVB* forms the outline contour of the upper lip, the contour curvature of which is the attribute *SVB\_curve*. The superior labium (the lip itself) varies in thickness and the de-

gree of protrusion (*SL\_fullness*, *SL\_protrusion*, and *SL\_thickness*). The total set of 16 mouth attributes summate to create the specific shape and dimensions of the mouth. While these attributes are all required to capture observed ethnic variations, further attributes could readily be defined to refine the model. Representative examples of combinations of the mouth attributes are shown in Figure 6.12.

The philtrum length *PHL\_length* (Figure 6.11) is noteworthy since it governs the Y position of the mouth as well as that of the chin below it. The philtrum, as defined here, extends from the subnasale (the bottom of the columella of the nose) downward, hence varying *PHL\_length* will result in translating the mouth and the chin in Y. Compare *PHL\_length\_min* and *PHL\_length\_max* in Figure 6.11, observing its effect on the placement of the mouth and its effect in Figure 6.12.

A broad variety of mouth shapes can be captured, as would be expected with 16 parameters. Importantly, when attempting to create an approximation to a particular mouth, the attributes act reasonably predictably, given that they are all summate to produce the overall shape of the mouth. Note that that the *position* of the mouth (in Y) is not an attribute

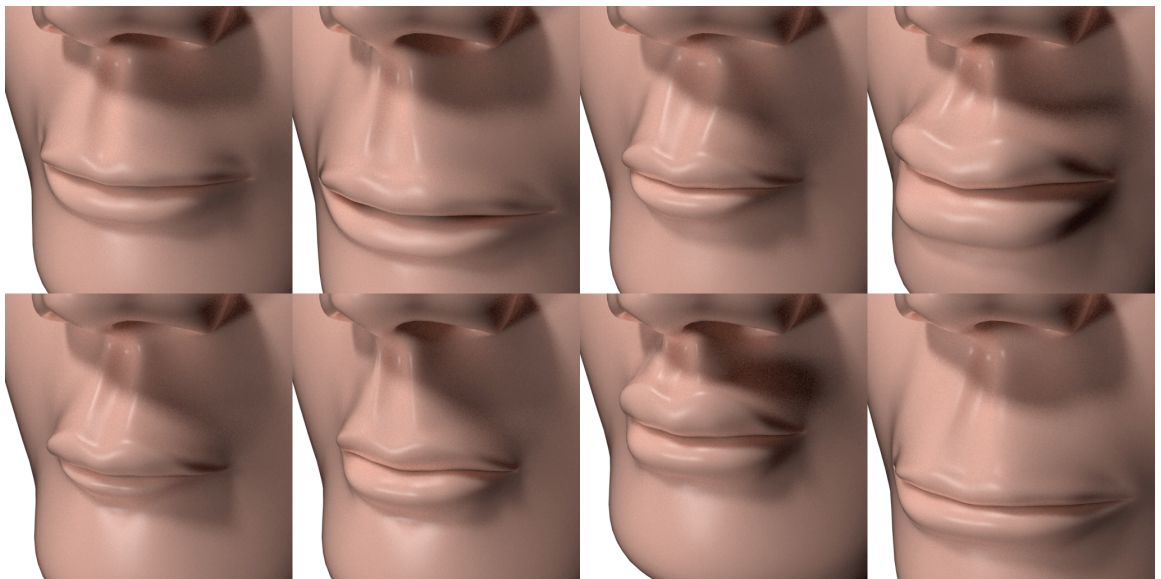


Figure 6.12. Eight examples of mouth shapes, some chosen to show extreme values for the attributes in order to demonstrate the range of resultant shapes possible.

in this system: it is a consequence of the length of the philtrum above it, which is, in turn, a consequence of the length of the nose above the philtrum, as discussed next.

### 6.4.3 Nose Attribute Basis Shapes

The nose region has six features (*ALA*, *COL*, *DSM*, *RAD*, *SID*, and *TIP*), which together contribute 17 attributes. These are comprised of four unsigned attributes (*ALA\_drop*, *COL\_drop*, *COL\_show*, and *TIP\_drop*), as shown in Figure 6.13.

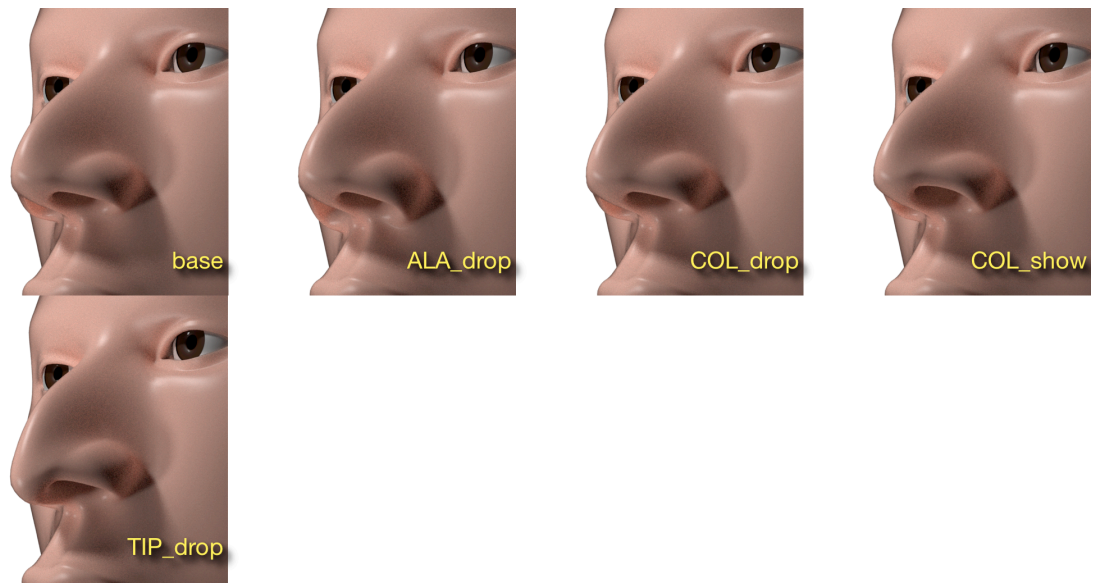


Figure 6.13. The base plus the four unsigned nose attributes are shown: *ALA\_drop* (the relatively low placement of the outmost margin of the ala), *COL\_drop* (the extent to which the columella hangs below the nostrils), *COL\_show* (the elevation of the nostrils to reveal the nasal septum in side view), and *TIP\_drop* (the relative displacement in Y of the very apex of the nose). The signed attributes are shown in Figures 6.14 and Figure 6.15. Refer to Table 5.3 and Figure 5.10.

The 13 signed attributes are shown in Figures 6.14 and 6.15. Some of the attributes of the nose determine its profile as seen in side view, while others determine the breadth of the nose along the ridge of the nose, as seen in front view. The nose is especially complicated to describe, since many attributes necessarily overlap spatially and yet need to combine without blendshape interference.





*Figure 6.14. Eight of the 13 signed attributes of the nose (ALA\_contour, ALA\_width, COL\_width, DSM\_length, DSM\_protrusion, DSM\_width, RAD\_protrusion, and RAD\_width) are each represented by their extremes. For example, the width of the nose varies between that represented by ALA\_width\_min and ALA\_width\_max. Refer to Figure 5.10.*

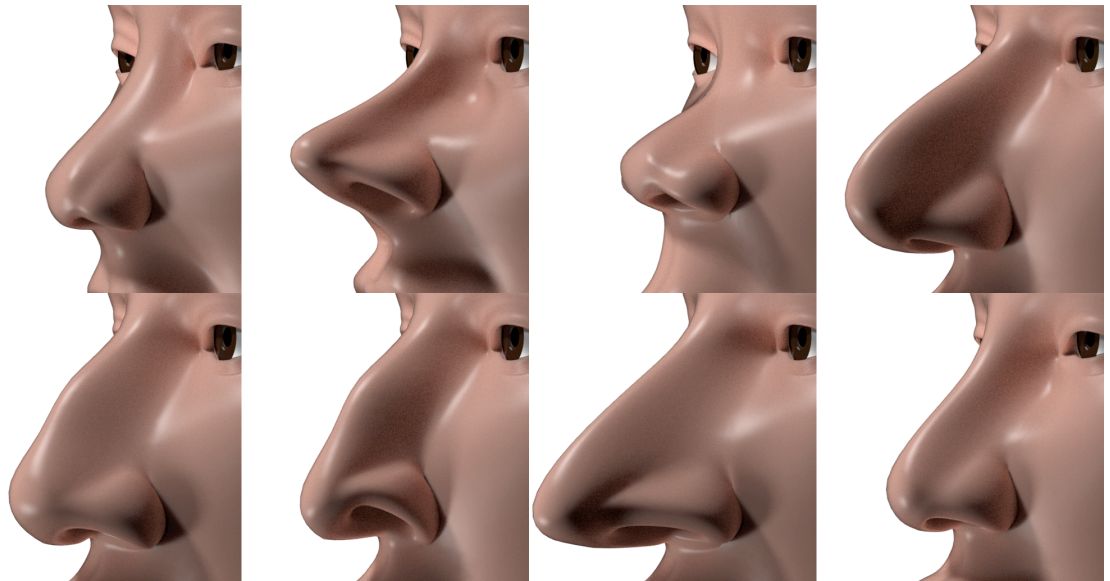
The distinctive lateral profile of an individual nose is the consequence of the relative protrusion at the radix (*RAD\_protrusion*), the dorsum *DSM\_height* (especially protrusion the rhinion at midpoint along the ridge), and at the tip, *TIP\_protrusion*, and the columella *COL\_drop* (see also Figure 6.15 and refer to Figure 5.10).



Figure 6.15. The basis shape pairs for the remaining five signed attributes of the nose: *SID\_slope*, *TIP\_inclination*, *TIP\_pointed*, *TIP\_protrusion*, and *TIP\_width*. See also Figure 6.13 and refer to Table 5.3 and Figure 5.10.

Since the protrusion attributes (from *RAD\_protrusion*, *DSM\_protrusion*, and *TIP\_protrusion*) modify the model mesh primarily in Z while the width attributes (*RAD\_width*, *DSM\_width*, etc.) modify the model mesh in X primarily, they function quite independently to create a large space of possible nose shapes, as shown in Figure 6.16. While the radix has attributes *RAD\_protrusion* and *RAD\_width*, it can also be given a further positional attribute corresponding to the location of the nasale in Y. This would allow adjusting the lowest point of the nasal bridge.

Recall that the model mesh is fixed at the origin, and the nose extends downward in Y from the nasale to the subnasale, depending upon the length of the dorsum *DSM\_length*. The philtrum then extends from the subnasale to CPB (Figure 5.7), for a length *PHL\_length*. Therefore the position of the mouth depends upon both the length of the



*Figure 6.16. Examples of combinations of nose attributes. The profile is governed primarily by protrusion attributes while the breadth of the nose and the slope of the sidewall are determined by attributes that are quite independent of the protrusion attributes. A very large space of possible combinations result, some clearly unlikely exaggerations, others quite representative of observed nose shapes.*

nose and the length of the philtrum. Mouth position is not an explicit attribute in this system.

#### **6.4.4 Overlapping Features**

Referring back to Figure 6.2, recall that the six facial regions were depicted with non-overlapping, sharply-defined borders (in either the polygonal or the subdivision surface representations), much as a geographical map delineates countries. The eye, mouth, and nose regions were then each described as a separate ‘module’ with features that were local to that region, despite the fact that they share common boundaries. Their attributes were then modelled by basis shapes that represent the extremes of each attribute (either signed or unsigned). Finally, it was shown that the features for each region could be assigned a wide range of combinations of attribute values to successfully create a wide variety of noses, mouths, or eyes. Little attention was given to the borders of any of these three regions, since most of their features lay completely within the boundaries of

that region. An exception was in the eye region: the interpupillary distance (*IPD\_x*) attribute did affect other features such as the palpebronasal sulcus (*PNS\_depth*) between the eye and the nose, the palpebrotemporal fossa (*PTF\_depth*) between the eye and the temples, and in the nose region, the width of the radix (*RAD\_width*), since the lateral shift of the eye necessarily compresses or expands the regions medial to, and lateral to, the eye. A satisfactory solution involved modelling the two extremes of *IPD\_x* by a pair of basis shapes, where the eyes as a whole were alternately shifted laterally to one or other extreme, then for each basis shape, slightly adjusting the area immediately adjacent to the repositioned eyes. Hence *IPD\_x* was represented by a single attribute even though that attribute affects adjacent regions.

In modelling facial attributes by basis shapes, the recurring goal is to achieve quasi-orthogonality, i.e., allowing the assignment of arbitrary combinations of attribute values across a set of features without adverse interactions. The “adverse interactions” in modelling *IPD\_x* were minor and easily finessed by small adjustments in the immediate vicinity of the eyes. More substantial issues, however, arose in modelling the nose.

In modelling the nose, the challenge was to define a set of attributes that, in different combination, could create different nose profile contours. The attributes in question were *RAD\_protrusion*, *DSM\_protrusion*, and *TIP\_protrusion*, i.e., the elevation of the ridge of the nose as it ascends from the radix through the dorsum to the tip. The relative protrusion of these three features (roughly corresponding to the sellion, rhinion, and apex) together define the profile contour of the ridge of the nose as seen in lateral view. Therefore, initially the three features were regarded as spatially-separate, non-overlapping segments of the ridge. The expectation was that the protrusion at each point along the ridge would then correspond to the corresponding anthropometric landmark. At the base of the nose, *RAD\_protrusion*, for example, would correspond to the elevation of the sellion (measured in the Z direction), and at the tip, *TIP\_protrusion* would correspond to the elevation at the apex (again, measured in the Z direction). But it was found that when these three features were modelled as spatially *non-overlapping*, they

proved insufficient to create the large variety of smooth profiles (such as a ‘Roman nose’ with the characteristically high rhinion or a straight ridge of various overall slopes). Instead, it was found to be preferable for these three features to smoothly combine along the ridge, i.e., modelling them with substantial spatial overlap, pairwise. That is, the radix and dorsum overlap, and the dorsum and tip overlap, but not the radix and tip. The attribute *TIP\_protrusion* was then modelled to affect the overall slope of the ridge as it ascends from the radix to the tip, and *DSM\_protrusion* changed the curvature of the mid-region of the ridge between concave to straight to convex, and finally *RAD\_protrusion* determined the height of the base of the nose. By adding the effects of the three attributes independently, different combinations permit modelling very different contour profiles, including the prominent rhinion (a ‘Roman nose’) and a supratip break, *even if those profiles are not explicit features of the model*. Interestingly, virtually all combinations of these attributes were plausibly found in the space of variation in the human nose. So rather than attempt to keep features separated and to finesse or minimise ‘adverse interactions’ (i.e., treating the results of some combinations of attribute values as introducing imperfections and therefore to be avoided), it was actually useful to design some features with substantial overlap and to allow their attributes to summate in a beneficial way. This shift in approach subsequently led to improvements in the design for the cranium, midface, and jaw regions.

#### 6.4.5 Cranium, Midface, and Jaw Attribute Basis Shapes

Compared to the eye, mouth and nose regions, the three remaining facial regions (the cranium, midface, and jaw) are more difficult to regard as separable modules. In retrospect, this can be predicted by how these regions are mapped in Figure 6.2: while the mouth, eyes, and nose are delineated by convex boundaries, the remaining regions of the face are concave. While some of their boundaries align with underlying osteological structures, such as the supraorbital ridge and the zygomas (cheekbones), other boundaries such as between the cranium and the midface (in the general location of the temples) and between the midface and the jaw are not so clearly demarcated. They nonetheless need to smoothly fair from one region to the other across their common boundary.



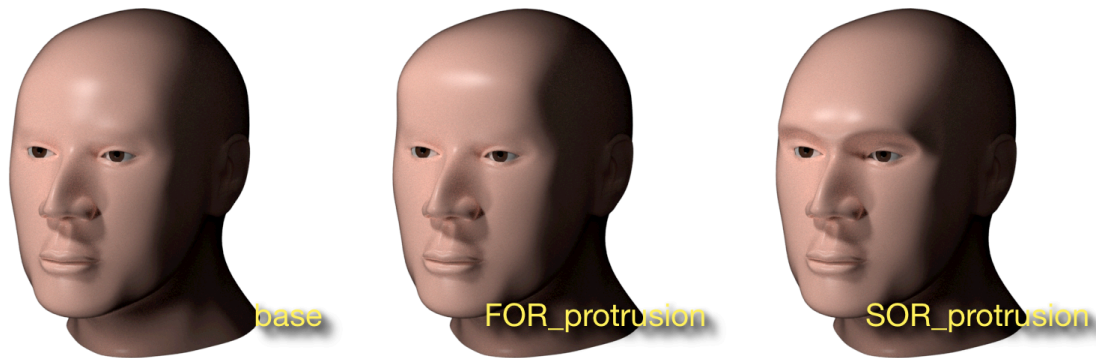


Figure 6.17. The base plus the two unsigned cranial attributes *FOR\_protrusion* and *SOR\_protrusion*. Refer to Table 5.4 and Figure 5.14.

Consider now the features and attributes of the cranium region. The cranium borders with the nose and eye regions along the supraorbital ridge (the horizontal protrusion of the frontal bone in Z). This Recall that the origin of the model is the midpoint between the left and right corneas (Section 6.3.3) and the eyes are fixed in Z and Y (but their IPD can vary in X, of course). All cranial features are relative to this origin. The 10 attributes of the cranium are shown in Figures 6.17 and 6.18. Two attributes are unsigned (Figure 6.17) and an additional eight are signed (Figure 6.18).

The supraorbital ridge (SOR), with attributes *SOR\_protrusion* and *SOR\_height*, determine the placement of the brow relative to the eyes. Only for very low SOR is there be an issue interaction with the features of the eye region by compressing the features of the superior palpebrum. Otherwise, the cranium region fairs with the nose region smoothly in the vicinity of the glabella, so that *RAD\_protrusion* permitted different types of nose profiles to blend into the forehead in a natural manner.

Next, the forehead attributes *FOR\_curvature*, *FOR\_protrusion*, *FOR\_slope*, and *FOR\_width* required many revisions, as these terms, while intuitive and seemingly independent, are actually tightly coupled. For example, changing *FOR\_protrusion* (in Z) necessarily changes the slope of the forehead, and changing *FOR\_slope* necessarily changes its protrusion. Both attributes are important in modelling different ethnicities, so



Figure 6.18. The eight signed cranial attributes *CRN\_depth*, *CRN\_height*, *CRN\_width*, *FOR\_curvature*, *FOR\_slope*, *FOR\_width*, *SOR\_height*, and *TMP\_width*. Refer to Table 5.4 and Figure 5.14.

it was decided to allow some interactions between the two attributes, requiring iteratively adjusting the two attributes to create a satisfactory match with a given forehead. The overall shape and dimensions of the cranium are governed by the attributes *CRN\_depth*, *CRN\_height*, and *CRN\_width*. Based on the osteology of the skull, the width attribute was also responsible for adjusting the X position of the tragon, and additional adjustment to the shape of the cranium in the temples just behind the orbits provided by adjusting *TMP\_width*.

Figure 6.19 shows a sampling of eight combinations of these cranium, forehead, supraorbital ridge, and temple attributes. Some examples are intentionally extreme to show the space of possible cranial shapes allowed by this set of 10 attributes.

Next, the midface and jaw regions incorporate attributes created by the underlying maxilla, zygoma, and mandible. The features and attributes of the midface and jaw were tabulated in Table 5.4 and labeled in Figure 5.14. While naming the features and their attributes was straightforward, creating three-dimensional representations of their extremes as basis shapes was surprisingly challenging.

Initial attempts at modelling the midface and jaw were met with the same problems as faced the modelling of the ridge of the nose, i.e. whether to regard their features as spatially-disjoint versus overlapping. Recall that the nose profile was modelled by treating the dorsum as superimposed upon the overall sloping ridge of the nose from radix to tip. The attribute *DSM\_protrusion* creates a perturbation at mid-length along the otherwise straight line contour of the nose as seen in lateral view. Analogously, when it

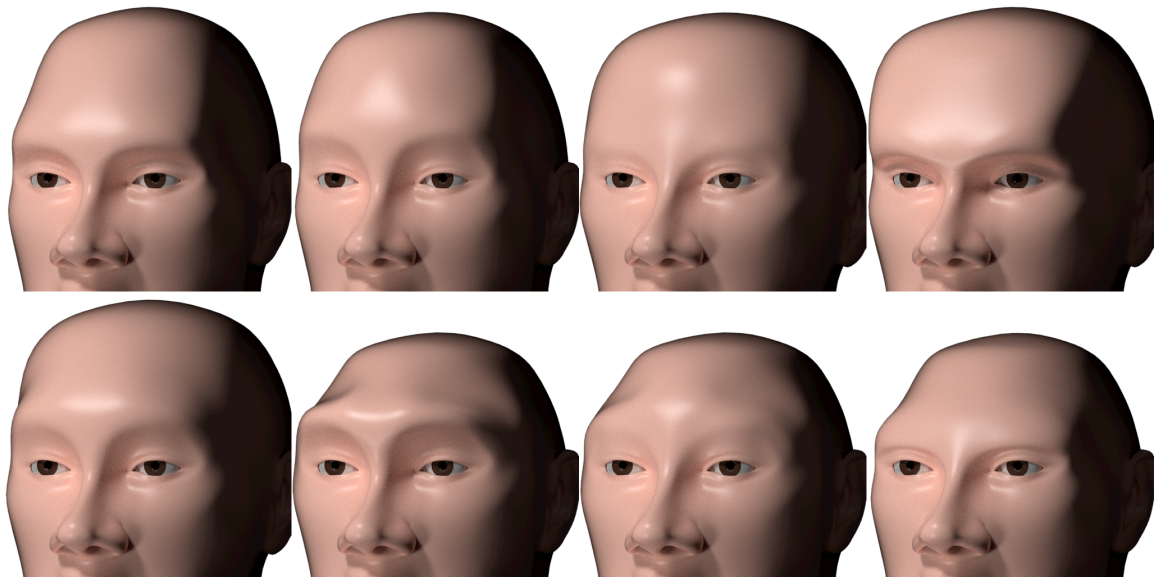
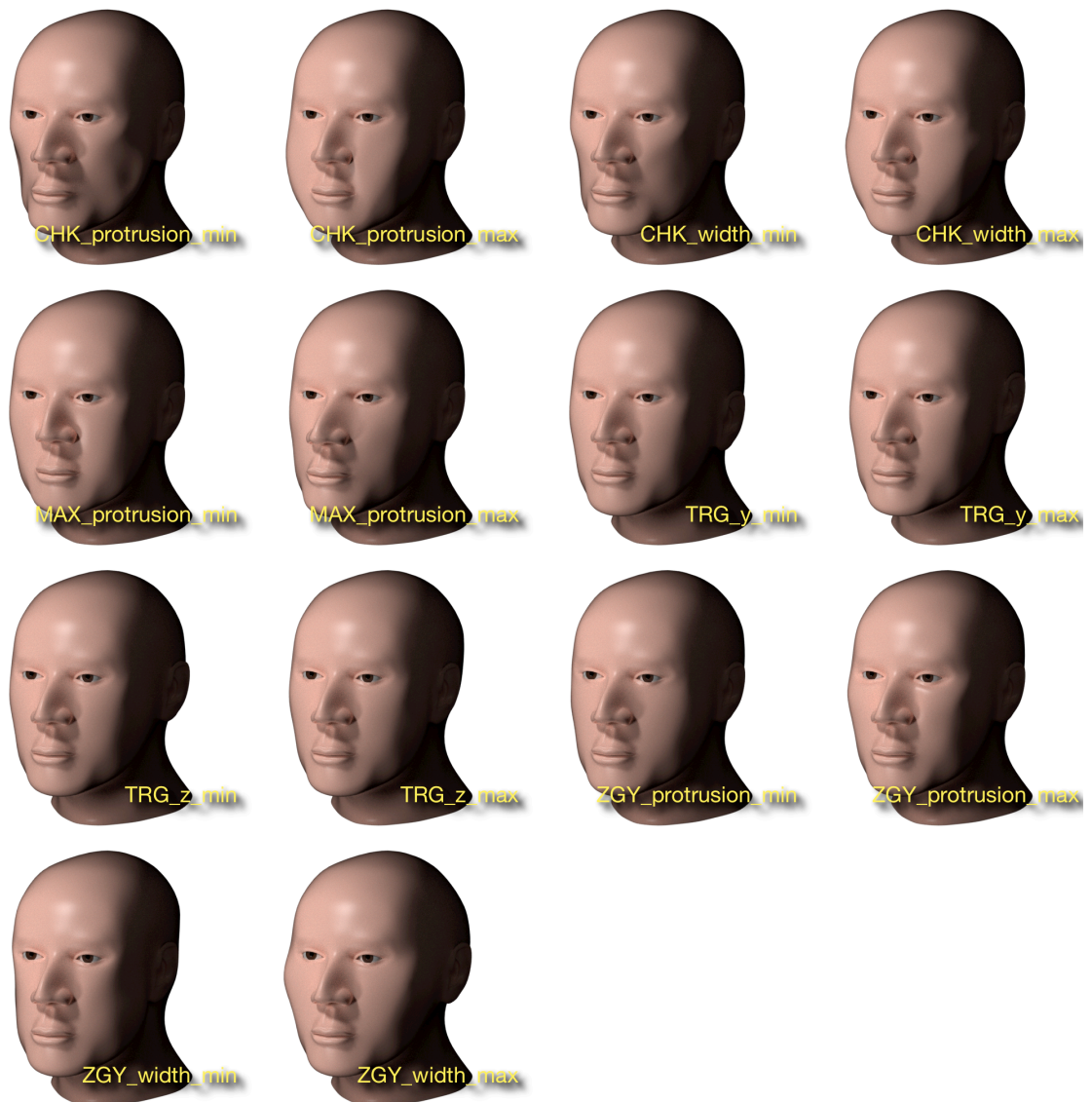


Figure 6.19. Various combinations of the cranium attributes (see Table 5.4 and Figure 5.14). These eight examples are chosen to demonstrate extremes of combinations of these attributes.





*Figure 6.20. The seven signed attributes of the mid-face region. The cheek (CHK\_protrusion and CHK\_width) and the zygion (ZYG\_protrusion and ZYG\_width) have substantial spatial overlap, permitting a variety of smooth facial shapes in that region. The protrusion of the maxilla (MAX\_protrusion) governs the overall placement of the nose, mouth, and jaw regions.*

comes to modelling the midface region, the entire nose and mouth regions are best regarded *superimposed* upon the maxilla (which they are, anatomically). Then, by adjusting the attribute *MAX\_protrusion*, the degree of nasal protrusion and prognathism in the mouth are controlled in a manner closely matched to the underlying osteology.

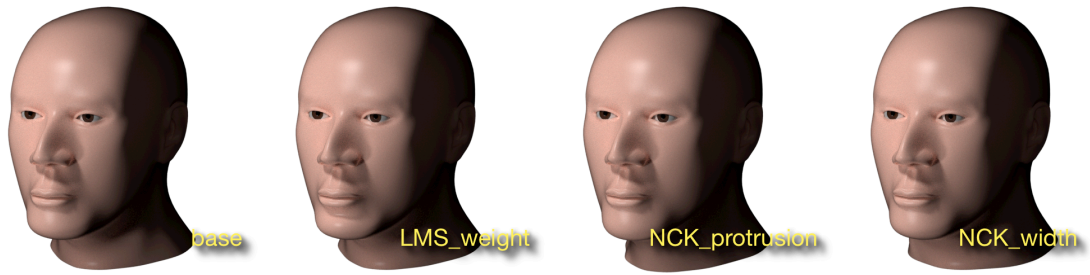


Figure 6.21. The base plus the three unsigned attributes of the jaw region: the weight of the labial mental sulcus (*LMS\_weight*) and the protrusion and width of the neck (*NCK\_protrusion* and *NCK\_width*).

More generally, the structure of the midface and jaw is well described in terms of either protrusion (in Z), width (in X), and elongations (in Y). Protrusions include *CHK\_protrusion*, *MAX\_protrusion*, *ZYG\_protrusion*, *GNA\_Z*, *MDB\_protrusion*, and *NCK\_protrusion*. Breadth attributes include *CHK\_width*, *ZYG\_width*, *GNA\_x*, *GON\_x*, *MDB\_x* and *NCK\_width*. The elongations or lengths include *GNA\_y*, and *GON\_y*, for example.

Next, the attributes of the midface region are shown in Figure 6.20. An important attribute is *MAX\_protrusion*, as this determines the protrusion of the nose and mouth (Section 5.4. and Section 6.5.3) as well as the jaw. This coupling reflect the osteology of the skull.

Lastly, the jaw region consists of the following features: the labial mental sulcus (LMS), the neck (NCK), the gnathion or chin (GNA), the angle of the jaw, or gonion (GON) and the mid-jaw or mandible (MDB). Figure 6.21 shows the three unsigned jaw attributes *LMS\_weight*, and *NCK\_protrusion* and *NCK\_width*.

Figure 6.22 shows the signed attributes are the placement of the gnathion (*GNA\_x*, *GNA\_y*, and *GNA\_z*), and gonion (*GON\_x* and *GON\_y*) and the protrusion and width of the mandible (*MDB\_protrusion* and *MDB\_width*). As discussed (Sections 6.4.2 and 6.4.3), the placement of the mouth is governed by the sum of dorsum length (*DSM\_length*) and philtrum length (*PHL\_length*) in Y, and in Z by the placement of the



*Figure 6.22. The seven signed attributes of the jaw region. This region has the following features: the chin or gnathion (GNA), the angle of the jaw or gonion (GON), and the mandible (MDB), with attributes governing the shape of the chin (GNA\_x, GNA\_y, and GNA\_z,) the placement of the angle of the jaw (GON\_x and GON\_y), and the protrusion and width of the mandible (MDB\_protrusion and MDB\_width).*

mouth in Z depends upon maxillary protrusion (*MAX\_protrusion*). In fact, the entire mouth region is further shifted in Z by the mandible protrusion *MAX\_protrusion*. This chain of displacements in Z is essential to capture the dependence of the protrusion of



Figure 6.23. Examples of various combinations of jaw and midface attribute values.

nose on the maxilla, and the protrusion of the mouth on both the maxilla and the mandible.

Finally, the gnathion is also dependent upon the maxilla, mandible and its own attribute `GNA_z`, all consistent with modelling the effects of the underlying osteology. Some representative combinations of jaw and mid-face attributes are shown in Figure 6.23. The midface and jaw attributes are primarily dimensional in nature, i.e., contributing primarily to the proportions.



## 6.5 The Ethnicity Modeller

The Ethnicity Modeller was developed in Maya as an experimental platform for exploring morphable models. The development process was iterative, and eventually a base mesh was created plus a set of copies with the same mesh topology that represent the attributes. The unsigned attributes requires a total of 22 blend targets (one for the extreme of each attribute) plus the set of signed attributes requires 55 additional pairs of targets (one for each of the extreme), resulting in a total of 132 copies of the base mesh.

The targets were assigned pairwise to blendshape nodes to represent each attribute, with all those cascading to influence the shape of the base mesh as diagrammed in Figure 6.3. This hierarchy underwent multiple, complete revisions to accommodate improvements to the topology of the base mesh (which then had to be ‘baked’ to all targets, a laborious process) and to accommodate the addition of new targets as new attributes needed to be added. Along with these large-scale revisions, countless small adjustments to individual targets were required which would often result in further adjustments to other targets to minimise unwanted interactions.

As the structure eventually became stable, the same base mesh as used for experimentation could then be used for rendering with the Arnold Renderer in Maya. The task of creating these 132 blendshape targets, while inefficient and time consuming, was consistent with a standard industry practice in character animation.

Rather than create a character rig, as conventionally used to control the application of blendshape weights, the set of target weights were under the control of a simple user interface written in Python that used Maya user interface widgets (popup windows, pulldown menus, sliders, etc.). The interface (see Appendix 1) permits a user to modify the value of any attribute, to view the mesh from any viewpoint, to introduce 3D photogrammetric data for matching and calibration, and to save a set of attribute values as a named model for subsequent loading and analysis.

The ethnicity modeller separates the data of a facial model from the polygonal meshes that represent the shape of the face. In the following sections the model data is discussed followed by the process by which the attribute values are adjusted to create a specific model.

### 6.5.1 The Model Data

The attribute data itself is economically stored in a file, as a set of dictionaries, one per attribute. Each dictionary consists of a list of multiple *key:value* pairs using the JSON (JavaScript Object Notation) format. Three string attributes are assigned: the attribute's name, the facial region, and a description text for documentation. In addition to the value, the range of the attribute is provided and, in some cases, the dimensions (in millimetres) of the extremes for that attribute are provided. For example, the dictionary for the attribute *ALA\_width* is:

```
{
    "description": "alar width al-al(X)",
    "attribute": "ALA_width",
    "region": "Nose",
    "maxValue": 1.0,
    "value": 0.53,
    "minValue": -1.0,
    "minDimension": 31.0,
    "MaxDimension": 50.0
}
```

The interpretation of the specific value 0.53 for *ALA\_width* is an interpolation between the associated basis shapes that represent the minimum and maximum extremes of *ALA\_width*. That is, this specific value of alar width corresponds to 53% of the range from the narrowest and widest nose. In order to compare *ALA\_width* with conventional anthropometric *al-al*, the minimum and maximum dimensions of the corresponding mesh were also measured (in this case, 31-51 mm for *al-al*). The attribute value 53% would then correspond to an *al-al* of 45.5 mm. A total of seven attributes have been identified as corresponding to dimensions from the anthropometric literature or readily measured

from images, and their range has been calibrated in terms of corresponding dimensions in millimetres. These attributes are alar width (*ALA\_width*, or *al-al*), cranium width (*CRN\_width*, or *t-t*), interpupillary distance (*IPD\_x*), zygoma width (*ZYG\_width*, or *zy-zy*), jaw width (*GON\_x* or *go-go*), mouth width (*CH\_width*, or *ch-ch*), and dorsum length (*DSM\_length*, or *n-sn*). These will be useful in model evaluation, including precision estimation and comparison with anthropometric literature (Chapter 7).

For this study, the model mesh is used to appreciate the range of variation of each attribute. Since basis shapes are implemented as polygonal meshes, they can also be exported to files in, for example, OBJ format (developed by Wavefront Technologies). Through the user interface (Appendix 1), a model can be read and its attribute values applied to the corresponding target weights in order to visualise the model, then the model maybe be saved after editing its attribute values (Appendix 2).

### 6.5.2 Matching the Model to Two-Dimensional Images

To adjust the model to match the shape and proportions of an actual face, initial experiments attempted to match the three-dimensional mesh to an ‘image plane’, i.e., a photograph of a face projected onto a plane in the three-dimensional scene (either an image of an individual or an averaged image). The focal length of the camera model in Maya was set to that of the camera that took the picture (from Exif data) and the camera model (through which the scene was viewed) was placed at the appropriate distance from the model. Looking at the model from this perspective, the two-dimensional face image could be seen as superimposed in the same three-dimensional space as the model mesh, which could then be adjusted to match the image plane face.

The process of two-dimensional calibration was ultimately abandoned due to several factors. First, the image averaging process in Webmorph resulted in images being normalised to constant *IPD*. This resulted in variations in all other dimensions accordingly, were which meant the images could only be used to model the shape of the features, not their dimensions or proportions (see Section 2.2.3 and Figure 2.3). Instead of matching

the model to an averaged image, an individual photograph was also introduced as an image plane. While this avoids the problems of normalisation, it revealed another issue — perspective distortion, wherein features such as the mouth are closer than the eyes and relatively too large due to perspective and features such as the ears are relatively too small. Even in this simple case, with a scale bar available in the photograph, the distortions were excessive, allowing only to match shape, not dimensions. More fundamentally, using a single two dimensional image as a template for a three-dimensional model is of course restricted because the model could only be viewed from the single fixed perspective at which the photograph was taken, and since the photograph collapses the face to a plane, only those facial measurements in that plane could be compared. Toggling the image plane on and off, or making it partly transparent could only provide a rough visualisation of the closeness of fit between the model and the photograph across the face. Photographs from multiple perspectives would have been useful, but of course that leads to a far better solution: stereophotogrammetry, the reconstruction of a three-dimensional representation from multiple viewpoints.

### 6.5.3 Matching the Model to Three-Dimensional Meshes

This study was also provided with photogrammetric three-dimensional face meshes of individuals and averages (Holzleitner et al., 2014), corresponding to both individuals and the computed mean of a sampling of twenty individuals of common sex and ethnicity (EUR or EAS). The ‘data meshes’ (as they will be referred to in order to distinguish them from the model mesh), were generated using the DI3D™ stereophotogrammetry surface imaging system (Winder et al., 2008), with delineation using MorphAnalyzer 2.4.0 (Tiddeman et al., 2000).

The model mesh is built around the placement of the two eyes (Section 6.3.3). Both the model and the two eyes are fixed but for one degree of freedom: the interpupillary distance *IPD*. A three-dimensional data mesh is then brought into registration by first aligning the corneas of the data mesh with those of the model, then adjusting the *IPD* of the model to match that of the data mesh. After aligning the corneas of the data mesh with



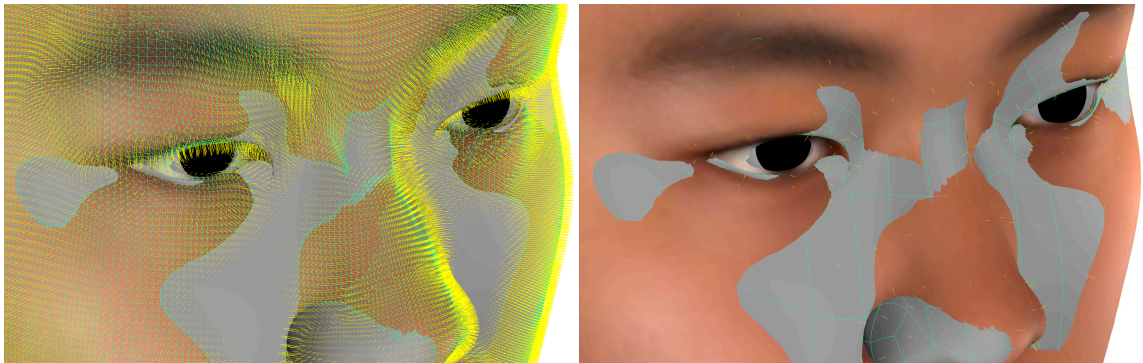
those of the model, one degree of freedom remains to be determined: the orientation of the Frankfort Plane for the data mesh. That is, the photogrammetric mesh might have a residual rotation about the X axis that would affect the placement of facial features, and that angular error would cause increasingly large disparity the further the distance of those features from the eyes. The protrusion of the chin, for example, would be modelled incorrectly if the Frankfort plane of the data mesh were not parallel to the XZ plane.

The Frankfort plane is defined by the uppermost point of the tragus and the orbitale, the lowest point on the margin of the orbit, (Swennen et al., 2006). Since the orbitale is not directly visible, it is estimated relative to the cornea. The mean eyeball diameter is 24.2 mm (transverse) by 23.7 mm (sagittal), with no significant difference with ethnicity or sex (Bekerman et al., 2014). The orbit dimensions are about 40 mm (transverse) by 35 mm (sagittal) (Ochs and Buckley, 1993; Sherman et al., 2006). The centre of the eyeball is a few mm below the centre of the orbitale, and therefore the orbitale is roughly 19 mm below the centre of the cornea (Stephan and Davidson, 2008). This disregards differences in orbit shape and dimensions associated with different ethnicities (Xing et al., 2013). With the location of the orbitale estimated, the Frankfort plane can be adjusted by rotating the data mesh in X until the line from the (inferred) orbitale to the tragus lies parallel to the XZ plane. This was required for a few of the data meshes that were provided.

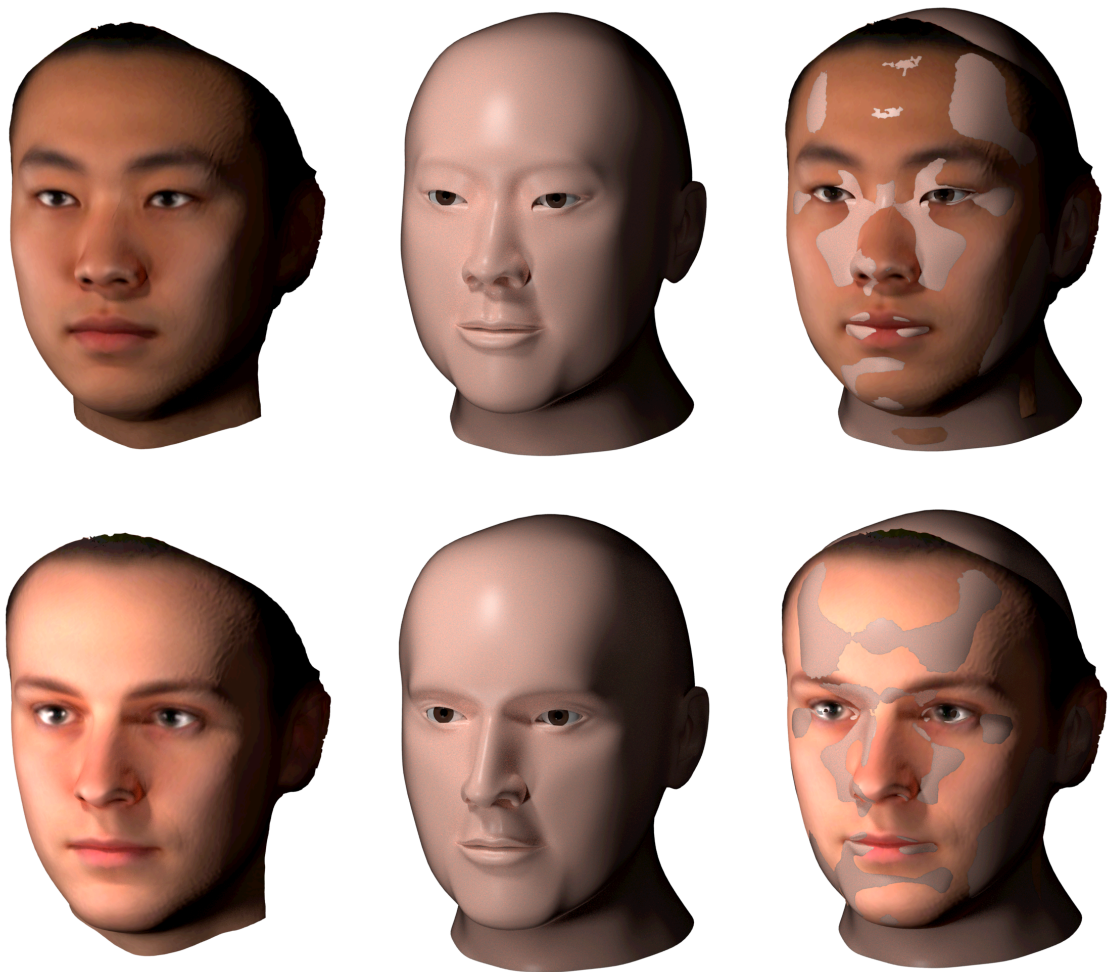
With a data mesh aligned such that the eyes were superimposed upon the model eyes and the Frankfort plane parallel to the XZ plane, the model was then adjusted through the user interface to bring the model mesh into close spatial registration with the data mesh. The first step is to focus upon the nasal alare (the point of greatest width of the nose), and to match this point in the model mesh with that in the data mesh, by adjustment of *DSM\_length* (in Y), *ALA\_width* (in X) and *MAX\_protrusion* (in Z). Just as the nose, mouth and chin are subject to the placement of the maxillary bone, the protrusion of the nose is relative to the protrusion of maxilla, as are the mouth, and chin.

After the protrusion of the maxilla determined, the attributes of the nose, face, jaw, cranium and eye can all be adjusted to match the data mesh in any order, however additivity of spatially overlapping features can permit the displacement of various regions to be achieved by various combinations of their attribute values. For example, the forehead has partly overlapping attributes *FOR\_curvature*, *FOR\_protrusion*, *FOR\_slope*, *FOR\_width*, and adjacent attributes *CRN\_height*, *SOR\_height*, and *SOR\_protrusion*. The process of matching the shape of a specific forehead requires successive refinement because since these features share multiple vertices, the displacements are additive, and a combination of attribute values must be found that summate to match the desired forehead shape.

The degree to which the model matches a given three-dimensional surface can be quantified and visualised by associating surface normals with one surface and observing to what extent those normals penetrate the other surface. Given two surfaces A and B, if A is just behind B, the normals of A will penetrate B, and vice versa. The exposed length of the penetration can then be used to measure the gap between the two surfaces.



*Figure 6.24. Modelling EAS\_M20, a data mesh representing the average of 20 male EAS. The closeness of fit is measured by 2 mm surface normals (yellow). On the left the normals are associated with the data mesh. On the right the normals are associated with the model mesh. By inspection the model mesh can be seen to fit generally well within  $\pm 2$  mm of the data mesh.*



*Figure 6.25. Stereo-photogrammetric meshes for the computed average of 20 male EAS (top left) and 20 male EUR (bottom left), their corresponding models (centre) and the two superimposed (right) to reveal their close similarity.*

In Figure 6.24 the closeness of the match between the model and the EAS\_M20 data mesh is demonstrated, where the normals are only 2 mm long, i.e., less than one standard deviation for anthropometric facial measurements within ethnotype, as discussed earlier (Section 2.2). In searching for where the fit is the poorest, note that the modelled side-wall of the nose, for example, is about 2 mm in front of the data mesh, and Figure 6.25 shows the data mesh for EAS\_M20 and EUR\_M20 (a stereo-photogrammetric average of 20 male EAS and 20 male EUR, respectively) in the upper left, the corresponding models in the middle, and the data meshes with the model superimposed on the right. When superimposed, the close match between the two surface meshes results in a patchwork of

interpenetrating surfaces, rather like the coloured patches of a Holstein or Guernsey cow (that pattern will be termed a ‘Guernsey pattern’ below). The patchwork is, of course, due to the two surfaces trading places on which is above the other.

In general, the modelling of facial attributes by linear-weighted basis shapes can successfully create a close match to a given data mesh, replicating the surface shape to within the variability observed for individuals with any ethnicity. The subtlety of features such as the nose, mouth, and eyes are well captured, but further refinement would be possible by the addition of further smaller-scale features and attributes, and subtle broader-scale general features to allow matching the fullness of cheeks, for example. The approach of modelling attributes through basis shapes is open-ended, and ultimately depends on the degree to which individual variations are to be captured.

## 7. Ethnicity Modeller Evaluation and Applications

The implementation of the *Ethnicity Modeller (EM)* provides a concrete test of the proposal set out in Chapters 3-6 for describing faces of differing ethnicities. Simply creating a sculpture that closely resembles a given face is not, in itself, a *description* of that face. But if that sculpture is based on a parametric modelling scheme (in essence a set of attribute-value pairs), the specific values of those parameters (interpreted with respect to the basis shapes) can be regarded as constituting a description. The fact that the same scheme can be used to model a very large range of faces provide the basis of a vocabulary or lexicon of primitive descriptors. Each face, whether the averaged face of a given ethnotype or the face an individual, would correspond to a specific set of attribute-value pairs, which together comprise a model.

The model is a quantitative description of a face which can also be visualised in three dimensions. The parameters are normalised and must be interpreted in terms of their associated exemplars (the specific 3D basis shapes that represent the extremes of each attribute). These exemplar basis shapes in effect anchor or create the foundation for the meaning of each attribute. The width of a mouth, for instance, could be expressed in terms of a numerical value and a pair of shapes that represent the extremes of mouth width. Provided those basis shapes capture mouth width in isolation from other attributes (such as lip thickness), then the concept of ‘mouth width’ has been made explicit, and furthermore, a specific mouth width is a specific interpolation between those extremes.

### 7.1 Evaluation

This schema may be evaluated in terms of several criteria. It consists of two aspects, the three-dimensional representation (the set of facial attributes and the corresponding basis shapes) and the user interface that controls the representation. The representation scheme provides the basis for describing faces, and would be evaluated in terms of three very concrete criteria:

- *precision* (how subtle of facial differences can be captured by this scheme?)

- *accuracy* (can the scheme create descriptions that match anthropometric measurements?)
- *orthogonality* (are the attributes capable of arbitrary combination)?

In addition, several other usability criteria that are more abstract, but nonetheless important:

- *intuitiveness* (are the facial attributes individually comprehensible and meaningful?)
- *efficiency* (does the interface permit practical assignment of each attribute?)
- *generality* (can a sufficient variety of facial types be represented?)
- *utility* (is the representation scheme useful for describing and distinguishing faces?)

The issues of representation (precision, accuracy and orthogonality) will be addressed first, followed by a discussion of the more abstract usability criteria.

### 7.1.1 Measurement Precision

In general, any measuring device has a limit in the precision of the measurements it provides, which is usually represented as an uncertainty interval around each measurement. In terms of conventional anthropometric measurements, for one individual, repeated measurement of the distance between two landmarks (such as *go-go*, *zy-zy*, or *ch-ch*) would result in a distribution of values, rather than precisely the same value each time. Differences in successive measurements would likely be caused by some combination of simple measurement error and imprecision in deciding on the location of the landmark, particularly for ‘soft tissue’ landmarks that are defined indirectly in terms of the underlying bone structure (Enciso et al., 2003). For example, the soft-tissue gonion (*go*) is defined as “... the most lateral point on the soft tissue contour of each mandibular angle, located at the same level as the 3-D hard tissue cephalometric gonion landmark” (Swennen, 2005). This landmark is particularly difficult to localise on the surface of the skin without reference to radiographic imagery (or direct palpation of the mandible), and even

then, there is uncertainty in locating “the most lateral point” on the underlying mandible. Since the distance *go-go* would be measured between two uncertain soft tissue locations corresponding to two uncertain osteological locations, error propagation would result in a greater expected measurement uncertainty for this distance than measurements that are based more directly upon well-defined superficial points, such as *en-en*, *ch-ch*, or *al-al*.

The absolute precision (e.g., in millimetres) that would be required of facial measurements (such as conventional anthropometric measurements between landmarks, Section 2.2) depends upon the specific application, of course. Anthropometric studies usually report statistical means compiled from multiple individuals, and therefore individual variation adds to the uncertainty of anthropometric measurements that are individually uncertain to measure in the first place. The statistical measurements reported in studies by Farkas (Farkas et al., 2005) and others (Section 2.2) reflect some combination of these multiple sources of uncertainty. Therefore, as discussed in Section 4.6, the *Ethnicity Modeller* need not match statistical data closer than the variance in that empirical data. The imprecision in anthropometric data suggests a practical limitation of several millimetres in any absolute measurement, in general. The numerical precision with which the *EM* can be used for anthropometric measurements was therefore examined.

Repeated measurements of any one anthropometric measurement of a given individual face would always be limited in precision, regardless whether performed with traditional physical callipers (Pérez-Pérez et al., 1990) or the digital equivalent applied to three-dimensional data. Likewise, the *EM*, when used to match a three-dimensional digital scan (of an individual or a digital composite of multiple individuals), would also result in measurements of limited precision. Section 6.5.3 discussed how the model mesh of the *EM* could be adjusted parametrically to match the surface from three-dimensional data (e.g., from photogrammetry). The surfaces of the model mesh and the data mesh could be brought into close proximity, i.e., mutually overlapping to within a few millimetres (Figure 6.24). This process involves considerable human interaction through a large

number of sliders, and many opportunities for measurement error during the matching task.

It was important to evaluate how precisely this matching process itself could be repeated. Repeated trials were performed in which the *EM* was initially in the default state, a data mesh (e.g., EAS\_M20) was superimposed in proper spatial alignment at the eyes and Frankfort plane, then all 77 parameters of the *EM* were iteratively adjusted until the model surface eventually matched the data surface sufficiently well, resulted (Figure 6.24). Obviously the procedure has many opportunities to introduce measurement error, which reduces repeatability.

The first step in examining the repeatability and measurement error quantitatively was to perform the surface matching process for multiple trials to produce a distribution of values for each attribute. The measured standard deviation of each attribute provides an estimate of measurement uncertainty. The details and results of the study are provided in Appendix 3.1.

As shown in Appendix 3, Table 3.1, the majority of attribute values showed small variances. Each standard deviation value was converted to a normalised fraction of the range for that attribute, providing a sense for the precision of the attribute. For example, the attribute *ALA\_width* had a measured mean of  $-0.39 \pm 0.17$  for EAS\_M20, and  $-0.55 \pm 0.17$  for EUR\_M20. Note that the EUR nose is narrower, as expected, but both have an uncertainty of  $\pm 0.17$ , which corresponds to roughly 5% of the range for this attribute. In other words, the uncertainty in an *ALA\_width* measurement is only about 5% in using the *EM* over repeated trials. This imprecision is comparable to the variability in statistical estimates based on conventional physical anthropometry that results from individual differences, as discussed (Section 4.6).

Similarly satisfactory estimates or measurement uncertainty were found for the majority of attributes (see green values for the columns “% EAS” and “% EUR” in Table A3.1).



But the repeatability was poorer for the more qualitative shape attributes such as *PM-S\_weight* (the palpebomalar sulcus, or hollow, between the lower eyelid and the cheek bone), *IPC\_weight* (the infrapalpebral convexity, or the fullness, of the lower eyelid), or *ECF\_weight* (the prominence of the epicanthal fold) are not so easily quantified. These shape attributes are unsigned, where 0.0 means non-existent (such as *ECF\_weight* = 0.0 for EUR, as they do not characteristically have an epicanthal fold), and 1.0 is an arbitrary definition of the maximum for that attribute (e.g., the most prominent ECF expected for any ethnicity). The just noticeable difference, or JND in these shape attributes is generally coarser than that of a dimensional attribute. In the repeatability study, all those measurements with standard deviations greater than 15% (and indicated by red in Table A3.1) were in fact shape attributes. While these shape attributes are not as precisely measured, they are nonetheless useful in modelling and describing the face. The precision with which shape attributes can be measured repeatedly is roughly half that with which dimensional attributes can be repeatedly measured. That is, the user interface might allow only 20 discrete steps in the sliders to match the roughly 5% JND for the dimensional attributes, and about half that many steps would be adequate to match the 10% or larger JND for the shape attributes.

The two composite data meshes used for this repeatability study will also be useful for measurement accuracy (Section 7.1.2) and ethnicity differences measurement and visualisation (Section 7.2).

### 7.1.2 Measurement Accuracy

To determine the absolute measurement precision (in millimetres) that can be achieved with the *EM*, an evaluation was performed for those attributes which have a direct anthropometric value. In the nose region, the attribute *ALA\_width* corresponds to the conventional measurement *al-al* and *DSM\_length* corresponds to *n-sn*. In the mouth region, *CH\_width* corresponds to *ch-ch*, in the mid-face region *ZYG\_x* corresponds to *zy-zy*, and in the jaw *GON\_x* corresponds to *go-go*.

These five attributes *ALA\_width*, *CH\_width*, *DSM\_length*, *GON\_x*, and *ZYG\_x* were then calibrated in terms of their dimensions in millimetres. For each attribute, the basis shapes representing their extremes were measured using the distance tool in Maya (which displays the distance between two locators). For example, *CH\_width* (the width of the mouth) was measured to vary between 40-54 mm corresponding to the range of -1.0 to 1.0. Any intermediate attribute value is then interpolated linearly between those extremes (a value of 0.0 for *CH\_width* equates to 47 mm, for example).

With these five attributes calibrated, the repeatability of measurements of these anthropometric measurements could then be examined quantitatively. Note that the majority of the attributes do not have such a direct quantitative relationship with traditional landmarks, either because conventional anthropometrics does not measure shape attributes (such as *ALA\_drop* or *SVB\_curve*) or because a conventional measurement would span multiple regions (such as the *n-gn*, the distance from the nasion to the gnathion, which comprises the dimensions of the nose, mouth, and chin). While *n-gn* is a traditional facial measurement, it is actually a composite of other measurements; the same overall value could be created by various combinations of a short-versus-long nose, a short-versus-long philtrum, thin-versus-thick lips, and short-versus-long chin. The *EM* only provides the local attributes, but through measurement tools the overall composite distance *n-gn* can also be measured directly from the model mesh.

The data from the earlier repeatability study was used to examine the accuracy of the five calibrated attributes to the corresponding anthropometric measurements from Farkas et al. (2005). The details are provided in Appendix 3.2. Corresponding to the generic composite of 20 EAS males (EAS\_M20) are two specific EAS ethnotypes, Singaporean Chinese (CN) and Japanese (JP). As shown in Table A3.2, the ethnicity modeller results are accurate within the published standard deviations of the anthropometric measurements *al-al*, *ch-ch*, and *n-sn*, with a difference that is less than one standard deviation. The match for the measurement *go-go* is also quite accurate for CN, but as the original

anthropometric data shows, the JP sample is about 10 mm wider than CN. This suggests that the EAS\_M20 data is comprised of primarily Chinese males.

To compare the EUR\_M20 data in to the anthropometric measurements for North American white male NA and German male DE, the results are again showing satisfactory accuracy for all but the *go-go* and *zy-zy* measurements. As discussed in Appendix A3.2, the gonion and zygion landmarks are based on features of the underlying osteology, and diagrammatic depictions place them in different facial locations by Farkas et al. (2005) versus Swennen (2005). It would appear that they are more useful for comparisons within a study (where the landmarks are presumably located consistently) than across studies (where the landmarks might be defined inconsistently).

In general, the potential has been demonstrated for the *Ethnicity Modeller* to provide absolute (millimetre) measurements for those dimensional attributes that are based on well-defined, superficial (rather than osteological) landmarks. The repeatability study suggests that the precision given the current user interface is, at best, on the order of a few millimetres, which is usable for measuring composite or average data given the variability across individuals, but unlikely of sufficient precision for measuring the model of an individual face.

It would appear feasible to extend the calibration process to other attributes, with two limitations. First, any attribute that is to be calibrated should create displacements in two landmarks that are well-defined by superficial features of the face, so that their locations can be reliably determined based inspection of the model mesh (thus avoiding the problems found earlier in calibrating *go-go* or *zy-zy* compared to the literature). Second, the distance between those landmarks should be influenced by only that one attribute. That is, there needs to be a 1-1 correspondence between attribute value and landmark separation for the calibration to be successful. By design, *ch-ch* and *al-al* are influenced only by one attribute each, namely *CH\_x* and *ALA\_width*, respectively. Other features, such as the breadth of the tip of the nose, unfortunately, are affected by multiple attributes and

therefore cannot be so directly calibrated. This is a matter of attribute orthogonality, which is evaluated further next.

### 7.1.3 Attribute Orthogonality

The 77 facial attributes were designed to be ‘quasi-orthogonal’, i.e., able to be combined in arbitrary combination, with a minimum of undesired complications. This requirement then permits modifying the width of the nose independently of other attributes of the nose. Most of the attributes of the *EM* are indeed quasi-orthogonal. First, there is isolation between the attributes in one region relative to the neighbouring regions (the eye attributes do not affect the nose attributes, and vice versa). Also, within the eye, nose, and mouth regions the attributes work ‘with a minimum of undesired complications’, however achieving that has been extraordinarily difficult and ad hoc.

*Attributes Dependent Upon Other Attributes:* In retrospect there is an aspect of the design, however, where orthogonality could not be adequately achieved, and where ‘undesired complications’ are probably unavoidable in principle. Recall that the only fixed point of the model is the origin — the midpoint of the imaginary line between the two corneas (Section 6.3.3). The entire face is then modelled by various combinations of attributes that shift localised regions of the face mesh relative.

Initially the model started with attempting complete independence between the nose, mouth, and jaw. While the goal was to create independence among the various attributes, it actually resulted in coupling. For example, increasing the length of the nose (by *DSM\_length*) resulted in shortening the philtrum (the mouth attribute *PHL\_length*). These coupling problems were resolved by revising the modelling of the nose, mouth, and jaw so their attributes were relative to *MAX\_protrusion*, the attribute which represents the protrusion (in Z) of the maxillary bone. In the human face, the maxilla is the foundation underlying the base of the nose and the upper tooth row, the soft tissues of the mouth, and the jaw and chin. The maxillary bone is, of course, not directly observed without radiography or tactile manipulation. It is nonetheless important as it influences

the forward projection of the mid and lower face. In the design of the *EM*, the basis shapes for the extremes of *MAX\_protrusion* were designed to shift the base of the nose, the mouth, and jaw, as does maxillary bone in life.

*Aligning the Alare:* Procedurally then, it was decided that the user would first align the nasal alare (the points of maximum width of the nose) in the model with the nasal alare in the sample data mesh (as derived by stereo-photogrammetry, for instance). The nasal alare were chosen because they are well-defined superficial points that are closely associated with the maxilla (the upper tooth row is just below the alare, as apparent by tactile examination). In the *EM*, to superimpose the alare in the model with those in the data mesh required carefully adjusting three attributes: *DSM\_length* (adjustments in Y), *ALA\_width* (adjustments in X), and *MAX\_protrusion* (adjustments in Z).

Once aligned, the modelling could proceed with adjusting the attributes of the nose and other regions. In fact, many other attributes of the nose, mouth and jaw were then critically dependent upon the initial choice of the three attributes *ALA\_width*, *DSM\_length*, and *MAX\_protrusion*. What is worrisome is that an error in *MAX\_protrusion* can be compensated for within limits by other attributes. For example, a slight underestimation of *MAX\_protrusion* can be compensated for by overestimation of *TIP\_protrusion* in the nose, *GNA\_z* of the chin and so forth. An error in *MAX\_protrusion* will become apparent only if it is so large that it prevents successful modelling of mouth or chin attributes. The dependency of several mouth and jaw attributes (those involving protrusions in Z) upon one attribute (*MAX\_protrusion*) is a significant limitation in achieving orthogonality.

*Testing for Blendshape Interference:* For attributes to be orthogonal, they must be able to be combined arbitrarily. Since they are implemented by blendshapes, they often suffered ‘blendshape interference’ (Section 6.3.1), a problem well-recognised in digital animation where it is common practice to control for undesired interactions by adding hidden ‘corrective blendshapes’ to counteract those problems. Unlike the solution in digital animation, it is not an option to repair attribute interference by behind-the-scenes blendshapes:

there must be a direct relationship between each attribute and the pair of basis shapes that represents the range of that attribute.

The first step in testing for orthogonality was to consider pairs of attributes that shared vertices within the same region. If these two attributes could be set to all four combinations of their two extremes, then another pair was selected, looking for cases that might reveal blendshape interference. If the various pairings of attributes did not reveal problems, then all attributes in the given region were set to various combinations of extremes (all minimum, all maximum, successive attributes alternating minimum and maximum, etc.) These ‘stress tests’ were not exhaustive, of course, due to the huge number of possible combinations, but the tests gave some confidence that some extremes of the space could be achieved without problems.

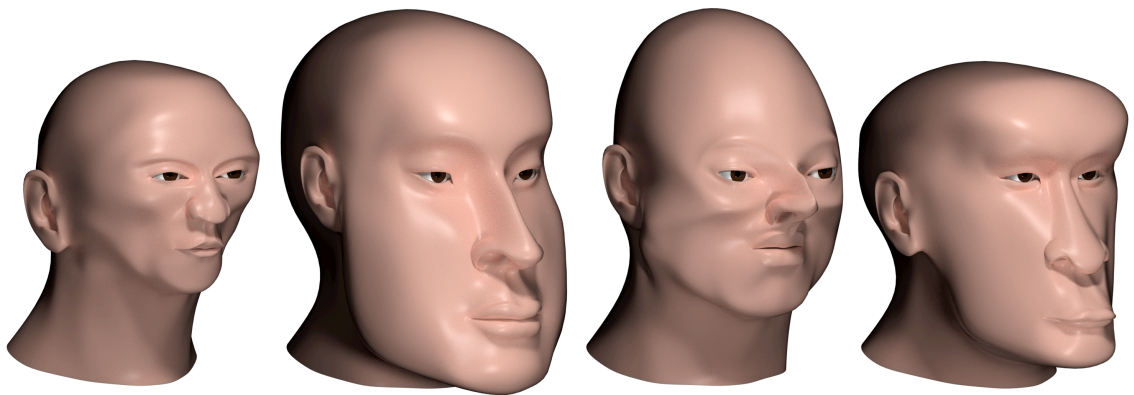
The 17 attributes of the nose region (after a considerable effort in adjusting their basis shapes) permit a very large number ( $2^{17}$  or 131,072) of combinations of extremes. This success is due, in part, to the natural partitioning of the nose into radix, dorsum, and tip ‘sub-regions’ and once the tip was given responsibility for the overall protrusion of the nose in Z and the dorsum given responsibility for the length in Z, then combinations of protrusion and length worked well together. Likewise the various width attributes (*RAD\_width*, *DSM\_width*, *TIP\_width*, and *ALA\_width*) work well to match a variety of human noses.

Figure 7.1 shows four extremes out of the  $2^{77}$  (i.e.,  $1.5 \times 10^{23}$ ) possible extremes of attributes. Clearly the face space is capable of representing an abundance of possible face shapes.

#### 7.1.4 Usability

In evaluating the schema the other, more abstract criteria concern usability: *intuitiveness*, *efficiency*, *generality*, and *utility*.

*Intuitiveness:* Intuitiveness, while highly subjective and abstract, was central to this approach from the beginning, starting with dimensional properties that were derived directly from conventional anthropometric measurements. Then, facial shape properties were then adopted from a variety of literatures, primarily medical, cosmetic, forensic, and physical anthropology. Of the many terms with which faces are described in these literatures, only those that refer to local shape features were used, as they are both intuitive and unambiguous, in contrast to global descriptors sometimes used to characterise overall facial shape or proportions (e.g., ‘oval-shaped face’).



*Figure 7.1. As a ‘stress test’ on attribute orthogonality, the models from left to right show all 77 attributes set to their minimum values, all to their maximum values, then alternating minimum then maximum, and alternating maximum then minimum. The two faces on the left are maximally separated in this face space (they lie in diagonally opposite ‘corners’ of a 77-dimensional cube) as are the pair on the right.*

Intuitively, local shape attributes such as the protrusion of the cheeks or the curvature of the forehead, clearly vary with ethnotypes. But they have not been previously quantified as distinct attributes. The *Ethnicity Modeller* demonstrates that these attributes can be quantified by interpolating between a pair of basis shapes representing their extremes. A given degree of protrusion of the cheek or curvature of the forehead can then be represented by a fractional value within those extremes, even if that fraction is not easily converted into a corresponding anthropometric measurement.

*Usability Evaluation:* Five unpaid evaluators participated in a usability study of the *Ethnicity Modeller* interface. Four were students personally recruited at the University of Oregon, in Eugene (three were undergraduates, SB, CK, and JK, and one a post-graduate, SS), and a fifth evaluator, JM, was a colleague, a senior professional with experience in computer-aided design and engineering in Portland, Oregon. All had high competence in using menu-driven user interfaces in general. Two students, SB and JK, had no prior experience with any three-dimensional software; all others had at least some experience (e.g., with Blender or Unity, and SS and JM had considerable skill in Autodesk software such as Maya and AutoCAD).

All evaluators were given the same three tasks. Individual sessions were performed with one-on-one supervision and interviewing, each requiring between 45 minutes to an hour. The interviews followed the ‘thinking aloud’ (Lewis, 1982) methodology. This methodology seemed best-suited to this project for revealing design issues, for which “...it is generally held that at least four to five users are necessary to detect the majority of the problems in a system” (Hertzum and Jacobsen, 2001); see also (Lewis, 1994). The tasks were intended to evaluate how well individuals could control the appearance of the model mesh through the use of the facial attribute sliders.

The first task was initially intended to use the EM to ‘sculpt’ a reasonable approximation to a face based on a two-dimensional image of EAS\_M20, with the very general instructions “*Please use the Ethnicity Modeller to create a three-dimensional model of this face. When you are satisfied with model, please turn to the next page to provide us with your feedback.*” This task (sculpting a model from an image) was replaced with a far more structured surface-matching task wherein the model was adjusted to match a 3D data mesh of EAS\_M20. This surface-matching task was reported as feasible by all evaluators, as will be discussed.

The second task involved presenting side-by-side images of EAS and EUR, and having the evaluator use the attributes as a vocabulary for describing differences in the facial



features of: *“Using the same facial attributes as your technical vocabulary, describe in words a few of the major differences between the East Asian and the European face.”*

The third task was to report on their experience using the interpolation slider tool of the *EM*: *“Please compare your experience with ‘morphing’ between the two faces in 3D with the alternative of shifting your attention between the two adjacent photographs.”*

In evaluation interviews with the five users of the *EM*, the adjustment of both shape and dimensional attributes were unanimously reported as intuitive even to novice users. While the specific terms such as ‘supraorbital ridge protrusion’ or ‘inferior palpebral convexity’ may have been unfamiliar, the underlying concepts were immediately apparent by manipulating the slider *SOR\_protrusion* or *IP\_convexity*. When the user was then assigned a facial region to model, each slider in turn was scrubbed it back and forth to see its effect on the model. Once the ‘meaning’ of a few sliders became familiar (i.e., their effect on the shape of the model mesh), the evaluators all expressed satisfaction that they had developed some facility in selecting combinations of their values that produced the desired shape. One user, JK, remarked that the shape adjustment process in the *EM* felt like he was a beginner using a colour picker to adjust a colour, first by trial and error while watching the effect that individual (RGB or HSV) sliders had on the sample colour swatch. For a given region such as the nose, he would experiment with one slider after another, using them with some success but not by memorising the names of the sliders as much as remembering their effect. User JM also remarked that if one does not know the technical meanings of the attributes one has to guess them by experimenting.

Another evaluator, SS, commented that while the attributes ‘made sense’ while being manipulated through the user interface, there was no definition for the spatial extent of each attribute. What, for instance is the extent of the radix of the nose, or the zygoma, or the gnathion? This question raised an important point, for while some features correspond to point-like landmarks (such as the cheilion) others do not have well-defined locations or boundaries. The map in Figure 6.2 is only schematic, since the vertices

along the boundaries of the facial regions are involved in multiple attributes and within any coloured region, several attributes may share many common vertices. Instead, some tool tip might be useful that annotates with diffuse boundaries the facial region affected by any attribute (similar to overlay provided by some tools in Adobe Photoshop).

*Efficiency:* While novice users may have found that individual sliders seemed intuitive effect, no user was satisfied with their matching results, and after adjusting a few facial regions, each user expressed discouragement with creating a close match between model mesh and data mesh. It was difficult to find a combination of slider values that would match the model to the data sufficiently well, especially in the forehead, mid-face and jaw. SS remarked *“I wouldn't say there were any attributes that seemed any easier or more difficult than one another as they all seemed to be weighted against other attributes on changing the value. This required constant adjustments when another slider seemed to put the wanted alignment in the desired place.”* Clearly the attribute sliders did not feel sufficiently orthogonal to SS. While SS was consciously aware of this issue, other naive users were probably frustrated by the same underlying problem.

The nose and eyes were easier to match, likely because the attributes were obviously orthogonal. In the nose region, for example, *DSM\_protrusion* (in Z) and *DSM\_width* (in X) were clearly independent, as were, in the eye region, the various attributes associated with the upper and lower eyelids. The cranium and midface, however, were much more difficult to model. While *CRN\_height*, *CRN\_depth*, and *CRN\_width* were obviously independent, that was not the case with *FOR\_slope*, *FOR\_protrusion*, *FOR\_curvature*, and *FOR\_width*. Similarly, the various attributes of the mid-face and jaw regions had broadly overlapping regions affected by multiple attributes in much the same way. For instance, *CHK\_protrusion* and *ZYG\_protrusion* both affect broad regions of the midface in Z, in order that each one blends smoothly into the other. Since *CHK\_width* and *ZYG\_width* also have some effect in the Z direction as well as their primary effect in the X direction, it became quite challenging to find some combination of those six attributes that matched the target. With the forehead, mid-face, and jaw there was a sense that one

could eventually get close to the desired target shape with sufficient ‘tinkering’, but there was little confidence about the choice of attribute values, since maybe a better combination might be found that would provide a better match.

The efficiency with which the *EM* can create a model is dependent upon the user interface as well as the skill of the user. Evaluation of the *EM* revealed that some aspects of the matching task are disappointingly inefficient, even for someone well familiar with the *EM*. Some novice users commented that there was no sense for when the match was sufficiently good, or how to make it better after a certain point. This sense of vagueness in creating a model is shared by those with much more experience with the interface.

The two novice users (SB and JK) independently commented that they wished they could just directly adjust the surface shape rather than using the sliders. They knew *what* they wanted to achieve, but felt it inefficient and frustrating to try to achieve it *indirectly* through the adjustment of sliders. While the attribute-value scheme would be preserved as the basis for describing faces, the user interface would not present the attributes for direct manipulation. This closely corresponds to a proposal by Lewis and Anjyo (2010) for the control of facial animation using blendshape weights that are hidden from the user. The animator would manipulate the face directly to form the desired expression while an inverse kinematics solver would determine the corresponding set of blendshape weights required to produce the desired manipulations. The benefit is clear: the user (as CK explained) is considering what they want to achieve, and not how to go about achieving it, just as an articulated figure can be more easily posed by moving limbs directly by inverse kinematics compared to the ‘forward kinematics’ adjustment of individual joint angles iteratively until the desired pose is achieved (Lewis and Anjyo, 2010). Adjusting an expression, or the shape of the nose, would seemingly benefit by hiding the implementation layer of blend shapes.

A similar suggestion to improve efficiency would be able to ‘pin’ or fix attributes once the user was satisfied with the shape in that region. This user was, in fact, reacting to the

fact that some attributes have undesired interactions due to the large degree of spatial overlap, such as *FOR\_slope*, *FOR\_protrusion*, and *FOR\_curvature*. Of course, if the attributes were truly orthogonal, that would not occur. Inter-dependence across some attributes frustrated several of the users that evaluated the *EM*. The solution is to patiently search for a ‘best fit’ combination of attribute values such as for the forehead, but this is non-trivial and very inefficient when attempted manually.

A further user comment regarding inefficiency concerned an attempt to using the *EM* as a sculpting tool. Rather than adjust the model mesh to some data mesh, the user was asked to start from the default face and to adjust attributes until the mesh resembled a face shown in front and side view as images. Modifying the shape of a face by adjusting sliders proved unintuitive to some novices, especially to individuals with little artistic skill.

An individual with significant experience in Autodesk Maya (SS) was especially useful in revealing problems with the *Ethnicity Modeller* when attempting to precisely match the surface of the model mesh with that of the data mesh. The data mesh, recall, that is created by stereo-photogrammetry consisted of a moderately low polygon count mesh on which an averaged image of skin shading and pigmentation was texture mapped. The resulting coarse, blurred quality of this averaged texture was difficult to localise in depth, and especially, it was difficult to critically perceive the relative distance between the smooth, featureless model mesh relative and the textured photogrammetric data mesh. Transparency would have helped, it was suggested, but counterintuitively, it was suggested to replace the blurred but supposedly ‘realistic looking’ skin texture map with a less realistic but more vividly three-dimensional geometric texture. Similar textures could be applied to both the data mesh and the model mesh to make it clearer which was nearer than the other. By further enhancing the sense of depth (for both the model and the data meshes) through stereoscopic presentation or motion (e.g., oscillating the viewpoint slightly to create parallax so that the relative separation in depth between one surface patch and the other was more vivid) could help the user adjust the attributes to bring one surface into alignment with the other. The display of the 2 mm surface normals that

show inter-penetration between the two surfaces was regarded as of only limited utility, and only in the immediate vicinity of the displayed normal.

A later study (see Section 7.2.2 and Appendix A3.4) in which stereo-photogrammetric data of individual subjects, added to the above observation. The stereo-photogrammetric texture for a individual (in contrast to the blurred, average of multiple individuals in the case of composites) was much sharper, and that provided a far clearer sense for the three-dimensionality of the surface. More broadly, the surface matching task is difficult, in fact more difficult than initially recognised, and improvements in the perception of the two surfaces in 3D should permit better matching of the model to the data.

*Generality:* The next overall quality of the representation scheme to consider is generality, i.e. the ability to model a sufficiently broad range of ethnic variation. Evaluation of the *Ethnicity Modeller* revealed two concerns about generality. First, there were persistent problems in matching some difficult patches of the face such as the top of the forehead along the midline, the width of the upper forehead compared to the lower forehead near the supraorbital ridge, the protrusion of the lower cheeks at the level of the mouth, the total distance from the top of the upper lip to the bottom of the lower, the profile of the gnathion as seen in lateral view, the location of the gonion in  $z$  (which would simply require introducing a  $GON\_z$  attribute), and the area below the temple between the eye and the tragus. These were to be expected, since the set of 77 attributes described above could easily become 100 attributes, or many more.

The second concern about generality was less expected: despite creating what appeared to be quite improbably exaggerated basis shapes to represent the extrema for various attributes, the experience with modelling individuals was that, in many cases, those attributes in the *Ethnicity Modeller* did not provide sufficient range of variation to adequately model some individuals (see Section 7.2.2).

*Utility:* Finally, it should be noted that the utility of the representation scheme underlying the *EM* is largely unproven. The evaluation of the *Ethnicity Modeller* has been limited thus far to commenting on precision, accuracy, intuitiveness, and generality (all of which show good promise) and efficiency (which clearly could benefit from improvements to the user interface and automation). Demonstrating the utility of the *EM* requires using the data delivered by the modeler for purposes of modelling individuals, ethnicity description, and static (caricature) and animated visualisation, as discussed next.

## 7.2 Applications

### 7.2.1 Quantitative Comparison

Given that two very distinct ethnotypes, EAS and EUR, had been modelled using the *Ethnicity Modeller* to examine attribute precision and accuracy, the data in those two models could also serve to quantify differences between the two ethnotypes, attribute by attribute. In Appendix 3.3, the pairwise differences are sorted by absolute value in Table A3.4 in decreasing order, so that the first few rows of the table correspond to the attributes with the largest numerical differences. The first few rows confirm what is already well documented in the literature: that EAS and EUR faces differ substantially in terms of the nose (*DSM\_protrusion*, *RAD\_protrusion*, *TIP\_inclination*, and *TIP\_protrusion*), eyes (especially the *ECF\_weight*), the cranium (*FOR\_protrusion*, *FOR\_slope*, *SOR\_height*, and *SOR\_protrusion*), and so forth for the jaw and midface regions. While reassuring, these differences are not un-expected: most of those attributes have already been identified as characteristics that distinguish the EUR versus EAS face. But since only a few of the nine attributes just listed have been measured quantitatively in conventional anthropological studies, it is promising to see that these attributes can now be successfully measured in the *Ethnicity Modeller*, and to a practical degree of precision. And beyond the obvious differences that distinguish between EAS and EUR, many novel attributes have been added to describe ethnic differences.

Given that the value for any specific attribute can be compared pairwise between EAS and its counterpart in EUR, can those differences themselves be rank ordered to reveal

which specific attributes show the greatest difference? Of all the 77 attributes, *RAD\_protrusion*, the extension of the radix of the nose in Z, turned out to have the greatest numerical difference between EAS (with a low value of  $-0.92 \pm 0.08$  quantifying the flat, low bridge of the nose characteristic of many EAS) and EUR (with the high value of  $0.59 \pm 0.11$ ). The difference between the two means was computed, with their uncertainties added according to error propagation, the difference in the attribute *RAD\_protrusion* was  $1.51 \pm 0.19$  between the two ethnotypes. This quantitative difference is not absolute (that is, the amount of protrusion of the nose in Z at the radix was not quantified in millimetres). Instead, this number only represents the protrusion relative to two basis shapes, one for the flattest radix expected to be observed in any ethnotype, and the other, the most protruding or tallest radix.

Calibration is straightforward for a few attributes, such as *al-al* and *ch-ch* (those five that are shown in bold in Table A3.4). A given difference for one attribute (such as *RAD\_protrusion* differing by  $1.51 \pm 0.19$  between EUR and EAS) is therefore not comparable with another difference, such as *ECF\_weight* (which differs by  $-0.79 \pm 0.04$ ). The eye of the EUR male characteristically has no ECF, of course. So a difference of  $-0.79$  in that attribute in EAS versus absence in EUR might in fact be far more visually salient in distinguishing EAS and EUR than a difference in the protrusion of the nose at the radix.

So while the overall table can be ordered according to the magnitude of the attribute difference, these differences are not strictly comparable across attributes. The magnitude of the difference for some attribute does not directly equate to a salience or significance in distinguishing ethnicities. Comparing *ECF\_weight* and *RAD\_protrusion* is a case of ‘apples and oranges’. Relatively large differences in attribute value across ethnotypes are nonetheless more salient than insignificantly different attribute values, however the specific rank ordering of differences is only a rough indication of this salience. At the bottom of the table, the smallest differences were less than 0.1 for many attributes (and smaller than their uncertainties). The two ethnicities clearly did not differ significantly in those attributes.

While a tabulation of attribute differences provides a rough comparison of two ethnicities, it is very difficult to appreciate these differences only by visual inspection of the numbers. For the five quantitative attributes that correspond to anthropometric dimensions (such as nose width *al-al*), the table is no better at helping to appreciate ethnic differences than other published tables in the literature. Of course, the goal of creating an extensive set of facial attributes is not to just produce and tabulate measurements; the goal is visualisation.

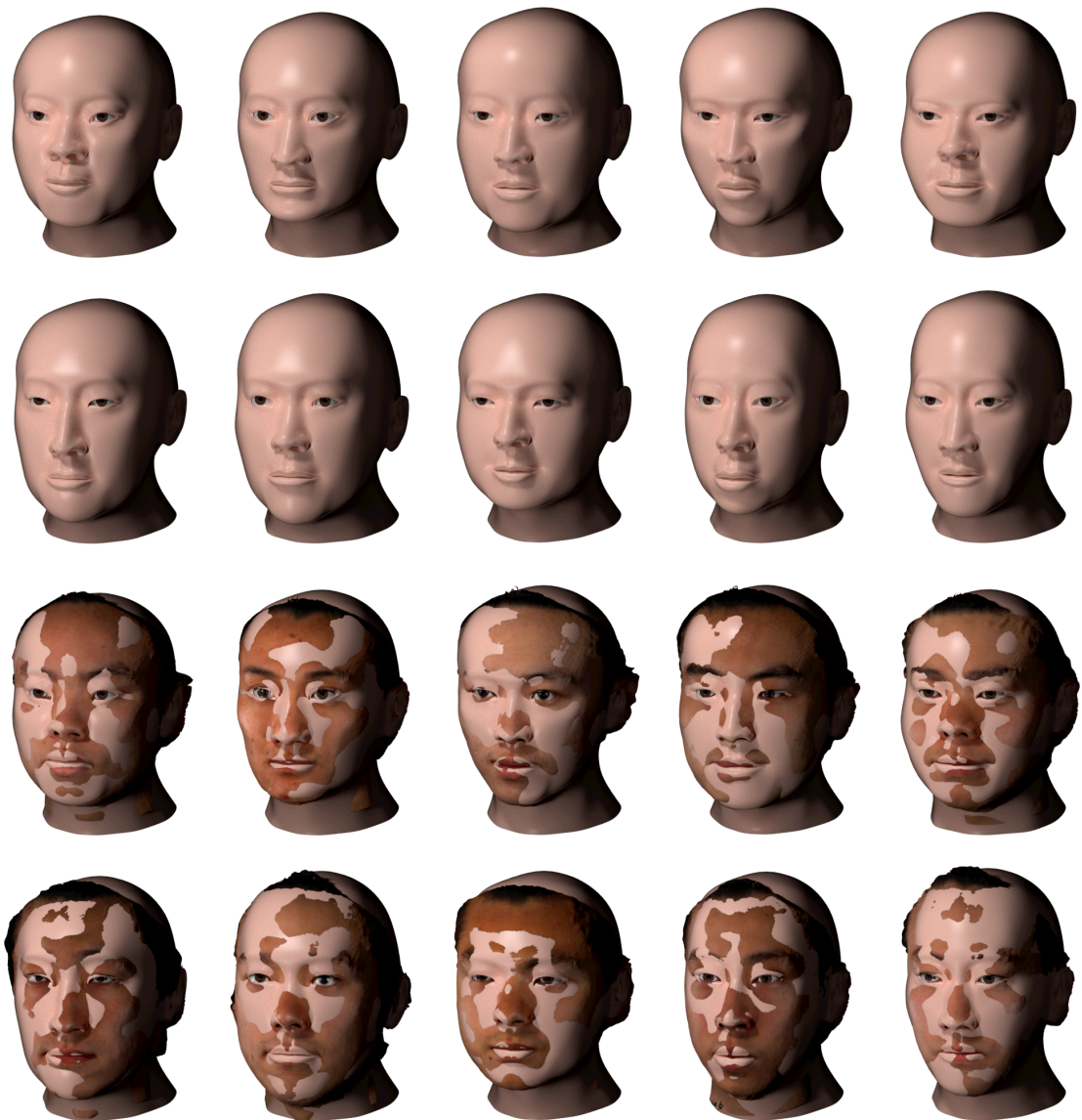
### 7.2.2 Measurement of Individuals

The repeatability, accuracy, and differences studies were based on two stereo-photogrammetric data meshes: EAS\_M20 is an average of the data meshes of 20 EAS male individuals and EUR\_M20 a composite of 20 EUR males (see Section 6.5.3). Fortunately, the original stereo-photogrammetric meshes were also provided for 10 of the 20 EAS individuals that comprise EAS\_M20 and for 10 of the 20 EUR individuals for EUR\_M20. These 20 individuals were then separately modelled using the *EM* to gain experience with fitting the facial attributes to individuals and to quantify individual variation within a population sample.

The natural asymmetry in individual faces (Farkas and Cheung, 1981) reduces the ability to create precise matches, but in most cases a compromise between the left and right sides provides a reasonable fit. The undistorted shape had to be estimated, which increased uncertainty in the adjustment of some attributes. An unexpected benefit of working with individual stereo-photogrammetric scans (rather than averaged composites of multiple individuals) was the greater clarity of the textures compared to those which are averaged composites. The data mesh was more clearly seen in depth and therefore more readily matched with the model mesh.

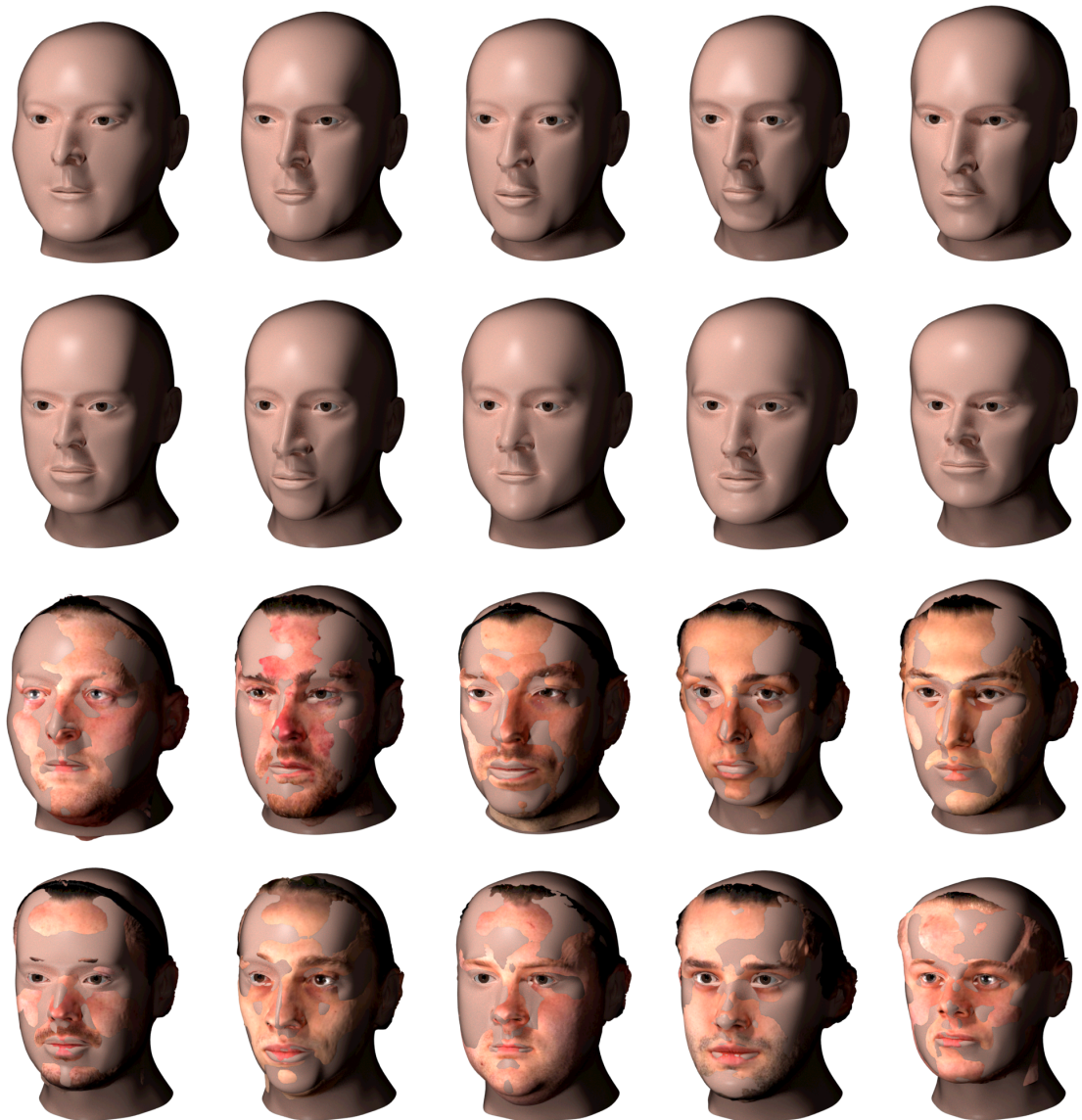
The 10 individual EAS models are shown in the upper two rows of Figure 7.2, and the lower two rows show those models in the same order but with the original data meshes superimposed. The resulting ‘Guernsey patterns’ of spatial overlap show the quality of





*Figure 7.2. The results of matching the Ethnicity Modeller to stereo-photogrammetric scans of 10 individual EAS males (top two rows), and superposition of scan data to show quality of fit.*

fit, and the left-right asymmetry in the patterns reflect irregularities in the individual faces (compare with the nearly symmetrical patterns in Figure 6.25). It was striking how much individual variation was present in these 10 EAS faces. That variability simply is not appreciated when the averaged EAS face is examined. Likewise the results of modelling the 10 individual EUR models are shown in Figure 7.3, along with overlays of the data meshes.



*Figure 7.3. The results of matching the Ethnicity Modeller to stereo-photogrammetric scans of 10 individual EUR males (top two rows), and superposition of scan data to show quality of fit.*

The 20 models (each a JSON file with 77 attribute-value pairs each) were then formatted into a spreadsheet in order to examine the distribution of values for each attribute across the 10 EAS individuals and the 10 EUR individuals. The means and standard deviations for the five attributes that have calibrations in millimetres (*ALA\_width*, *CH\_x*, *GON\_x*, *DSM\_length*, and *ZYG\_width*) were compared with the corresponding anthropometric data from (Farkas et al., 2005) in a follow-on to the accuracy study (see Section 7.1.2 and Appendix 3.2). The *Ethnicity Modeller* replicated the physical anthropometric measure-

ments well, with disparities of only a few millimetres (see Appendix A3.4, Table A3.5 for EAS and Table A3.6 for EUR).

### 7.2.3 Visualising Ethnic Differences

Parametric digital modelling of a face allows one to create a reasonable replica of the shape of a face, such as EAS or EUR. The model can be simplified to exclude skin tones, hair, skin texture, and so forth, to just concentrate on the shape of the face.

The appreciation of two shapes A and B by side-by-side comparison is difficult, as it requires shifting visual attention from one to the other shape, seeking out places where the shapes differ then to understand in what way they differ in each location (Section 2.3.3). When the two shapes are complex (as are human faces) it is particularly difficult to locate and appreciate the more subtle differences glancing from one to the other. Instead of side-by-side comparison, it is very effective for one three-dimensional model to continuously morph between the two alternative shapes A and B (and back to A, repeatedly). The task of visually shifting gaze back and forth in search for differences is replaced with the easier task of just watching places where the shape is morphing and to study those changes. The natural tendency for visual motion to attract attention makes this process of finding differences quite automatic. This is particularly useful when several changes occur simultaneously across the surface as the shape morphs between A and B.

*Two-Dimensional Interpolation:* Morphing, or linear interpolation, between two images is demonstrated in Figure 7.4 (see Section 2.2.3 regarding image averaging and interpolation). Two images were used as the ‘extremes’: the average of 13 EAS males (far left), and the average of 20 EUR males (far right). The intermediate three images represent 0.25, 0.5, and 0.75 blends between these two extremes. Of course, these five frames would be part of a longer sequence with finer interpolation steps to provide a motion sequence lasting a second or two. The many differences between these two ethnotypes become quite apparent in the motion sequence, and far more effectively than



Figure 7.4. A sequence of blends that interpolate between two averaged images, representing EAS (left) and EUR (right) males, with intermediate values of 0.25, 0.5, and 0.75. See also (Wisetchat et al., 2018).

can be appreciated in side-by-side comparison of static images. Unfortunately, if the images that comprise a morphing sequence (such as those in Figure 7.4) are scaled to constant inter-pupillary separation IPD, the dimensions of the face are not preserved through the sequence (see Section 6.5.2).

*Three-Dimensional Interpolation:* The *Ethnicity Modeller* performs parametric morphing in three dimensions between two selected models. Linear interpolation between a pair of models is achieved by interpolating each attribute between the values assigned to each model. All attributes, including  $IPD_x$ , are free to change as the mesh interpolates from one model to another (Wisetchat, 2018).

Suppose the value for some attribute  $A$  for model 1 is  $A_1$  and  $A_2$  for model 2. To compute an interpolate between models 1 and 2, attribute  $A$  is given a variable value  $A_v$  such that  $A_1 \leq A_v \leq A_2$ . Given an interpolation index  $i$  ( $0.0 \leq i \leq 1.0$ ), then  $A_v = A_1 + i*(A_2 - A_1)$ . Applying this interpolation to all 77 attributes would create a parametric blend or interpolation between the two models.

Parametric interpolation is demonstrated in Figure 7.5. In the upper row a male AFR model is morphed into EUR\_M20, and in the bottom that AFR is interpolated into EAS\_M20. In each case the computed midpoint between the two models serves as a fulcrum for comparison. Shifting to the ‘left’ (as  $i$  reduces towards 0.0) the face becomes increasingly AFR.



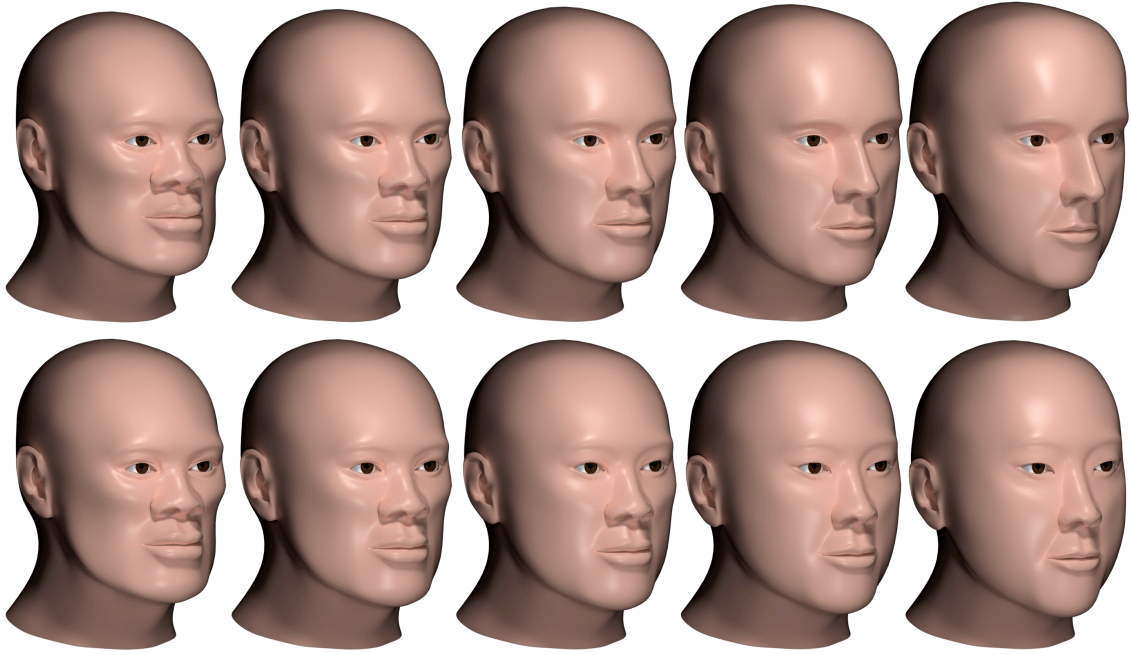
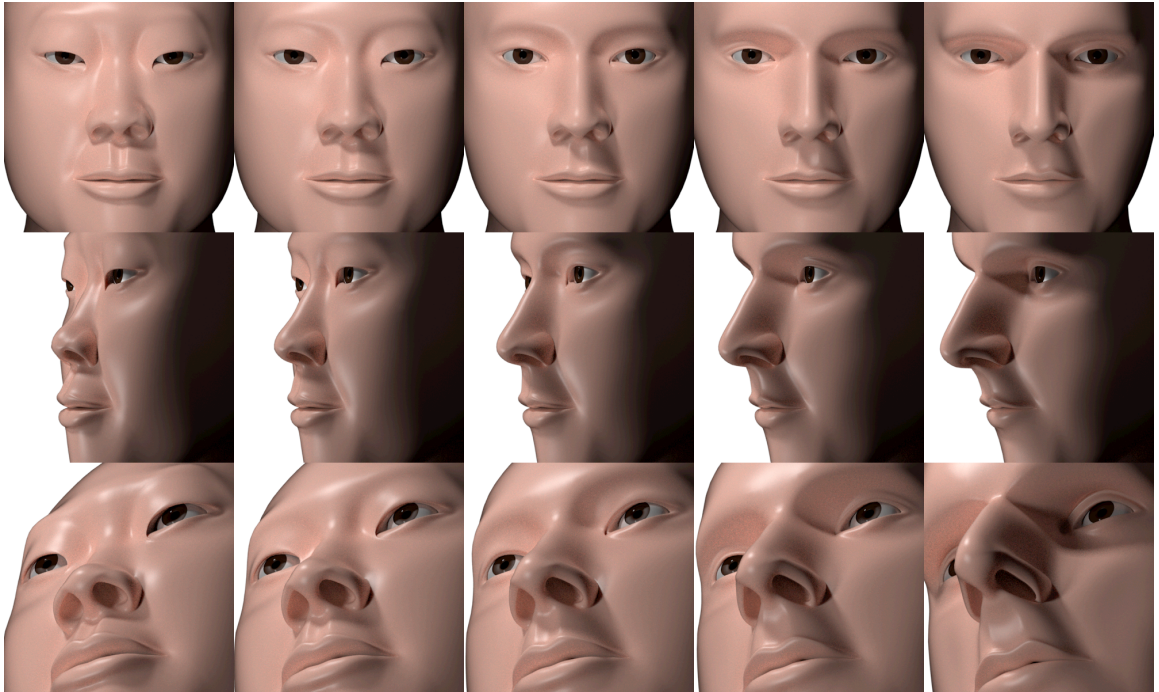


Figure 7.5. The top row shows a blend from an AFR model to EUR\_M20, and the bottom row shows that AFR model morphing into EAS\_M20. In both cases the middle image is a 0.5 blend. In each case the images reflect interpolations from  $i = 0.0$  (left) to  $i = 1.0$  (right). See also ([Wisetchat, 2018](#)).

*Emphasising Ethnic Differences through Extrapolation:* Note that interpolation becomes extrapolation for  $i < 0.0$  (creating an exaggeration of model 1) and  $i > 1.0$  (an exaggeration of model 2). Thus the *EM* could support caricatures (McIntyre et al., 2013), as shown next.

What would it mean to exaggerate the characteristics of an EAS face? Exaggeration is relative to some standard. Since, as discussed (Section 2.1.3), there is no global mean relative to which EAS features would be extrapolated, the exaggeration will be relative to *another* ethnotype.

Figure 7.6 demonstrates extrapolation EAS relative to EUR. The middle column of images show the mid-point interpolation ( $i = 0.5$ ) between EAS and EUR, which serves



*Figure 7.6. Creating a caricature by extrapolation. The three images in the middle column show the  $i=0.5$  interpolation between EAS and EUR from different perspectives. The columns to either side are the EAS ( $i = 0.0$ ) and EUR ( $i = 1.0$ ) models, and the extreme left and right columns show an exaggeration of EAS (left,  $i = -0.5$ ) and of EUR (right,  $i = 1.5$ ) relative to the mean ( $i = 0.0$ ).*

as the origin or fulcrum for extrapolation. The three rows show three camera views. Just to the left of the middle column is EAS\_M20 ( $i = 0.0$ ) and just to the right is EUR\_M20 model ( $i = 1.0$ ). On the far left the EAS characteristics are exaggerated by 100% relative to the mean for  $i = -0.5$ , and on the far right the EUR characteristics are exaggerated by 100% for  $i = 1.5$ .

The use of exaggeration (especially when viewed as an animation rather than statically) creates a very compelling demonstration of how these two ethnotypes differ (Wisetchat 2018). Some EUR individuals exhibit characteristics that are indeed an exaggeration of the mean EUR face, especially in the protrusion of a narrow nose, the low, jutting supraorbital ridges just above deep-set eyes, and narrow, hollow cheeks, as shown in the rightmost column of Figure 7.6. Likewise, some EAS individuals exhibit exaggerations

of the EAS ethnotype (shown in the leftmost column of Figure 7.6), such as puffy upper eyelids, heavy epicanthal folds and inclined palpebral fissures, a nose with low protrusion at the tip, considerable columellar show and tip inclination, and a face with low maxillary protrusion, yet prominent, wide cheeks.

The *Ethnicity Modeller* is capable of considerable exaggeration. In Figure 7.7 the upper row shows EAS (left), plus a 100% exaggeration (middle), and a 200% exaggeration (right). The lower row shows (from left to right): EUR, a 100% exaggeration, and a 200% exaggeration. The caricature process makes their mutual differences increasingly obvious even in these still images. With animated interpolation these distinctions can be appreciated without resorting to the distractions of exaggeration, i.e., by simply morphing between the upper left (EAS) and lower left (EUR) faces.

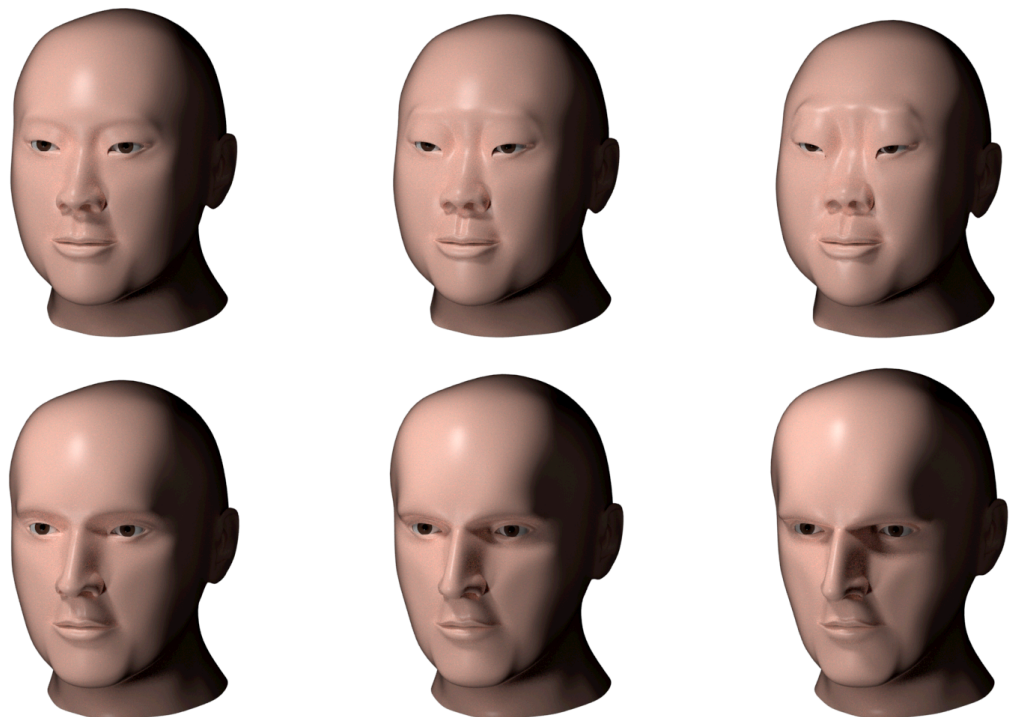


Figure 7.7. *Progressive extrapolation. The facial characteristics of EAS (top left) and EUR (bottom left) are subject to 100% exaggeration (middle) and to 200% exaggeration (right). See also ([Wisetchat, 2018](#)).*

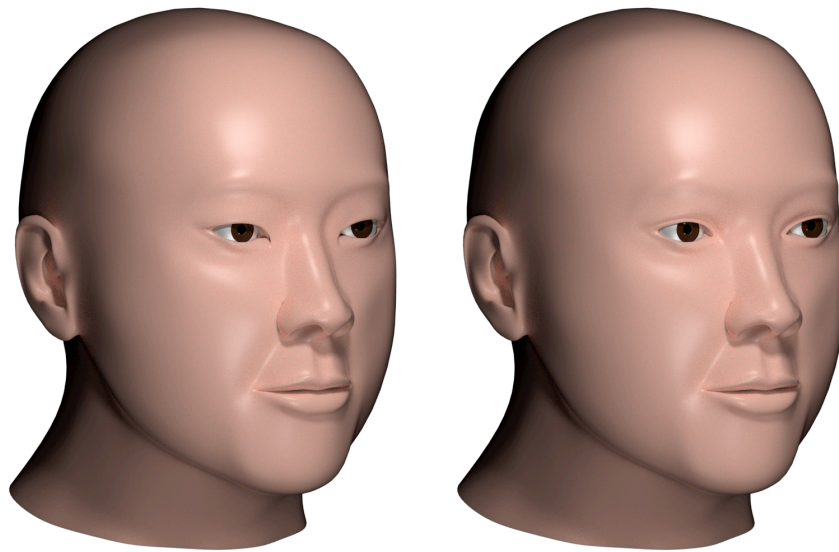
In terms of navigating a face space, linear interpolation between A and B corresponds to traveling along a straight-line path between A and B. Extrapolation corresponds to following that same path then continuing beyond B (to accentuate the characteristics of B), or in the other direction, proceeding from B to A then continuing farther to accentuate the characteristics of A. Extrapolation is provided by the *Ethnicity Modeller* between any pair of models (see interface controls in Appendix 1, Figure A1.3). One slider (the ‘all’ slider) controls all 77 attributes simultaneously, creating the interpolation and extrapolation in Figures 7.5 through 7.7. Travel through an ethnicity face space is then constrained to the straight line path between A and B, for any choice of the pair of models A and B.

All five users were enthusiastic that animated interpolation using the interpolation tool in the *EM* was an effective means for visualising ethnic differences. One individual, SS, commented on the value of creating ‘hybrid’ models (i.e., interpolates between two ethnotypes) as a way to see novel mixes of features, as well as the dynamic effect of seeing features in one ethnotype morph into those of another. The visual experience of watching animated interpolation was clearly preferred over any lengthly written description.

*Blending by Region:* The *Ethnicity Modeller* can be used to study the role of each facial region on ethnicity recognition, as demonstrated in Figures 7.8 and 7.9. Instead of interpolating all attributes in parallel with the same value, six separate interpolation controls (one for the attributes of each facial region) can combine independent blends of the six regions.

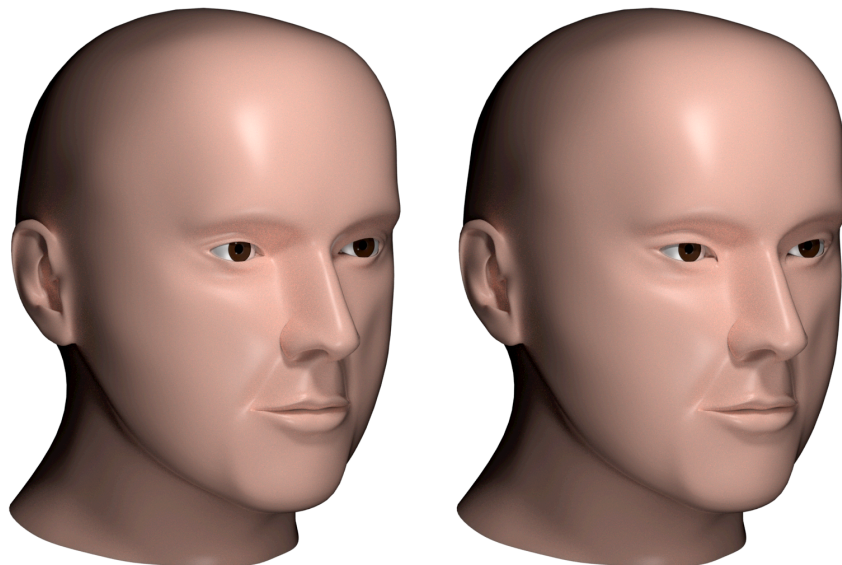
Figure 7.8 shows the Asian model EAS\_M20 on the left and the same model on the right except the eyes are now EUR rather than EAS. The preponderance of evidence still suggests that the face is Asian despite the inconsistent European eyes. Adding further EUR traits (to the nose especially) creates an increasingly European-looking face.





*Figure 7.8. The model of an average of 20 male EAS (EAS\_M20) on the left. The same Asian model is shown on the right but the eyes are EUR (from EUR\_M20) rather than EAS eyes.*

Similarly, Figure 7.9 shows a European face (EUR\_M20) on the left, and the same model is shown on the right except with EAS (from EAS\_M20) eyes rather than EUR eyes.



*Figure 7.9. The model of an average of 20 male EUR (EUR\_M20) on the left, and the same model but with EAS eyes substituted, on the right. The eye attributes are from the average of 20 EAS (EAS\_M20, see Figure 7.8 left).*

The EAS eyes appear inconsistent with the otherwise EUR facial features. This is but one possible mixture of region-by-region substitutions of ethnic characteristics.

### 7.3 Summary of the Evaluation

The evaluation process revealed the following about the *Ethnicity Modeller* interface:

1. *Measurement precision:* In repeated trials of matching the *EM* model mesh with the same data mesh, the resulting attribute values had standard deviations of generally less than 5% of the range of that attribute.
2. *Measurement accuracy:* Five attributes (*ALA\_width*, *CH\_width*, *DSM\_length*, *GON\_x*, and *ZYG\_x*) corresponding to anthropometric measurements (*al-al*, *ch-ch*, *n-sn*, *go-go*, *zy-zy*) were calibrated. Comparison of these attribute values (in millimetres) with corresponding anthropometric measurements shows close agreement, i.e., to within one standard deviation of their corresponding anthropometric measurements.
3. *Attribute orthogonality:* While all facial attributes were intended to be orthogonal (or quasi-orthogonal), and many do successfully permit independent adjustment, some dependencies are unavoidable due to the fact that some facial features are relative to others, which means their attributes will necessarily be coupled. A primary example is the dependency of nose and mouth attributes on the protrusion of the maxilla (*MAX\_protrusion*), corresponding to the fact that the nose and mouth are in fact built upon the maxillary bone. The benefit is that the nose and mouth are described relative to the underlying structure of the skull, but errors in estimating *MAX\_protrusion* will propagate into errors in those attributes that are relative to *MAX\_protrusion*. Another, less satisfactory failure of orthogonality is presented by attributes that overlap considerably within a common region, such as the forehead attributes (e.g., *FOR\_slope* and *FOR\_protrusion*) for which a combination can be found to successfully model the forehead, but which are not independent. Despite these issues, tests for blendshape interference show that the *EM* is capable of

successfully creating an enormous range of faces by random combination of attribute values.

4. *Intuitiveness*: The proposal to describe faces in terms of a common set of dimensional and shape attributes provides a novel framework for intuitively describing faces. A ‘model’ (i.e., the set of attribute-value pairs associated with a given face) is comprised of individually intuitive (or ‘semantic’) descriptors which are local (rather than holistic), and directly comparable across face models. The use of a pair of basis shapes to represent the range of each attribute is also both intuitive and novel as the basis for an ethnic face space (but of course standard practice in facial animation using blendshape deformers). The attribute set is extensible, and does not require assuming the existence of any ‘average’ face.
5. *Usability*: The *EM* interface requires expert familiarity with the terminology of facial features in order to select and modify the attributes by name. Novice users found difficulty in managing the complexity of the modelling task even within an individual region; as expected, the acceptability and efficiency of the *EM* depends on the skill and competence of the user. Future applications would be tailored to specific tasks; the *EM* was built only to directly control the exploration of attribute combinations.
6. *Generality*: The attribute set eventually consisted of 77 attributes, largely driven by attempts to closely model different facial regions. In combination, they provide sufficient generality to match facial features in data meshes, however the total set could easily be extended to 100 or more attributes. Unexpectedly, the generality of the attribute set was not in capturing sufficient detail so much as capturing the range of variation found in individuals. The basis shapes that represent the extremes of various attributes had to be modified numerous times as increasingly extreme examples were encountered.
7. *Utility*: The utility of the *Ethnicity Modeller* was demonstrated by various applications, including the modelling of individuals, the modelling of averaged examples of different ethnicities, the quantification of ethnic differences

in terms of their mean attribute values, model calibration, an economical representation for models (in terms of JSON-format dictionaries of attribute-value pairs), the ability to visualise ethnic differences by interpolating between models (effectively linear-path navigation within an ethnicity face space), the ability to create animated interpolations between models that control for which attributes are interpolated, the ability to extrapolate differences between two models in order to create caricatures, the ability to create randomised variations on a given model to produce characters, and more.

## 8. Discussion

An approach towards *describing* and *visualising* ethnic variation in the human face was developed. It consists of two parts:

1. a schema for describing a face in terms of a set of facial attributes (properties of discrete facial features), and
2. an experimental implementation of the schema called the *Ethnicity Modeller* which supports three-dimensional visualisation.

In retrospect, the schema introduced in this study follows the standard convention of distinguishing discrete features (in this case a total of 43 facial features were identified). Moreover, the 77 attributes — the properties of these features — include and expand upon a standard lexicon especially those that relate to anthropometric measurements, such as *ALA\_width*, which corresponds to *al-al*.

What differs from the conventional practice is the method by which attributes are represented. Rather than use informal descriptive terms such as ‘bulbous’, or discrete categories such as ‘platyrrhine’, or simply anthropometric measurements, this study uses a quantitative scheme wherein an attribute is represented by an interpolate between two extremes. By this means, the shape of a ‘bulbous nose’ can be represented by the combined attributes (*TIP\_inclination*, *TIP\_pointed*, *TIP\_protrusion*, and *TIP\_width*). Similarly, the term ‘platyrrhine’ can be represented, but it involves the combined effect of several nose features and their attributes (especially *TIP\_protrusion* and *ALA\_width*). The critical and novel addition is that the attributes *explicitly* represent the *shapes* corresponding to adjectives such as bulbous or platyrrhine. Each attribute is represented by a three-dimensional mesh corresponding to the corresponding quality.

The accuracy and precision of the representational framework was demonstrated and the practicality and usability of the user interface was tested through a series of studies involving stereo-photogrammetric scans of human faces. In retrospect, the results were mixed, and this chapter will critique these results, and in particular, attempt to explain

which aspects of the schema seem be clearly useful and worth pursuing, which aspects show potential, but for one or another reason have yet to be clearly demonstrated, and finally, which are less-than-successful (or at least very problematic) aspects that sounded very promising at the beginning, but might in fact revealing intrinsic limits to this schema.

## 8.1 About Face Space

### 8.1.1 An EFS Is Not an IFS

It is worth one final reminder of what this study is *not* about. The *Ethnicity Modeller* is *not* a scheme for automated face recognition (Section 2.4.1). Face recognition uses measurements that distinguish an individual from the mean of a population. The measurements are not the sort that are easily understood or visualised, as they are computed co-variation coefficients (such as those derived by PCA). Those dimensions, while orthogonal, and efficiently computed, and computationally useful, are meaningless to a human observer. Those dimensions form a multi-dimensional identity face space (IFS) which is conventionally limited to representing individuals from a single, homogeneous population, since the coordinates by which an individual is mapped into IFS is based on how much that individual differs from the average of that population.

Some aspects of automated face recognition that *are* incorporated in the present study, however. First is the recognition that faces vary along many dimensions, regardless which specific properties are chosen as dimensions. So it is useful to conceive of a ethnicity face space, but such an EFS is necessarily very different from the IFS, because the goal is to appreciate how faces vary across ethnicities (Section 2.4.2). Since there is no presumed average face across all ethnicities, the origin of the EFS is not important, but instead, the ranges of each dimension (Section 2.1.3). Most of those dimensions are not calibrated (in millimetres) but rather represent the extremes of various sorts of shape variation. For example, the shape of the tip of the nose varies from rounded to pointed. While this attribute varies with ethnicity, it is not expected to be one of the most salient differences; it is just one of the 77 attributes that were identified as salient. In fact, at-

tempts to distinguish salience (which components are ‘more principal’ than others) is not attempted in this study.

While cranio-facial landmarks provide a formal framework for *measuring* faces, face *description* is usually far less formal, and far less quantitative. A few anthropometric ratios (or ‘indices’) have been named and used as part of a simple descriptive scheme (e.g., to distinguish broad, flat noses from narrow, protruding noses, Section 2.3.2). The novel approach here is to develop a far more detailed, and hopefully far more useful descriptive language, one intended to be understood by humans.

The selection of facial features, and their attributes, does not have the mathematical rigour of PCA. Instead, they are ad hoc and intuitive, matching the properties that are recognised by studies of the face in various disciplines in addition to the conventional anthropometric measurements. In contrast to informal descriptive terms (such as simply labelling a nose ‘platyrrhine’), the facial attributes identified in this study are intended to also serve as EFS dimensions, i.e., the basis for exploring a whole space of ethnic variation.

### 8.1.2 Attributes as EFS Dimensions

Recall that a face is described by a set of attribute-value pairs, where each attribute is an EFS dimension and each value is a coordinate value along that dimension (Sections 2.4.2 and 4.11). In this study, the EFS has 77 dimensions, and while that is obviously too many to be navigated easily, each dimension is meaningful (or ‘semantic’). For this scheme to work, however, the attributes need to be: 1) continuous-valued, 2) capable of sufficient precision to permit reconstruction of a three-dimensional face to within some specific tolerances, 3) of sufficient number to capture the important aspects of facial variation, and finally, 4) orthogonal, i.e., able to be combined arbitrarily without ‘blend-shape interference’ (Section 6.3.1).

The first requirement, that the attribute values are continuous, is satisfied because they correspond to blend weights. The second requirement, providing sufficient precision, was satisfactorily demonstrated by the ability of the model mesh in the *EM* to match a data mesh generally to within 2 mm, i.e., with precision better than the variability due to individual differences (Section 7.2.1). Accuracy was also demonstrated: five dimensional attributes (calibrated in millimetres) were found to be in very close agreement with independently-measured anthropometric measurements (Section 7.2.1). Regarding the third requirement, that sufficiently-many attributes are defined, the conclusion is that there are *not*. Each region presented its own frustrations because there were always some details of shape that could not be matched with the available attributes. However, since additional attributes could easily be introduced (each with an associated pair of blend-shapes representing their extrema) this was not regarded as a significant defect in the schema; it just meant the 77 attributes were not quite sufficient — but an adequate number is probably not more than twice that. Moreover, when modelling individual faces it was often found that the extremes for various attributes had to be increased (exaggerated) to handle the surprisingly large range of individual variation of each attribute. The more extreme the differences between basis shapes, however, the greater the potential for blendshape interference, which brings us to the last requirement: orthogonality. Orthogonality was tested by examining combinations of attribute extremes, understanding that these represent the far ‘corners’ of the face space (Section 7.1.3) while modelling would generally stay well within the bounds of that space. As demonstrated in Figure 7.1, after much adjustment the attributes combine without much ‘blendshape interference’. But reflecting back on the process, persistent interference between two or more attributes was actually caused by the fact that those attributes are redundant, sometimes almost synonyms. For example *DSM\_width* (the breadth of the nose at the rhinion), *DSM\_protrusion* (the protrusion of the nose in Z at the rhinion), and *SID\_slope* (the slope of the sidewalls of the nose as they descend from the rhinion to the cheeks) are very inter-dependent: if *DSM\_protrusion* stays constant, increasing the breadth of the nose midway down the bridge would be expected to increase the slope, and decreasing the protrusion decreases the slope, etc. This necessarily causes blendshape interference, because the



three attributes simply are not orthogonal in principle. The attribute *SID\_slope* could have been omitted, but it serves a useful purpose in describing how quickly the side of the nose falls away, allowing EUR noses to be both narrow and steep, for example. So it was useful to have all three attributes even though they interacted. Likewise, some forehead attributes (*FOR\_curvature*, *FOR\_protrusion*, *FOR\_slope*, *FOR\_width*, *SOR\_protrusion*, and *SOR\_height*) were all more inter-dependent than initially expected when the modelling of the attribute basis shapes were begun in Chapter 6.

So some attributes, while intuitive, are not orthogonal, or even ‘quasi-orthogonal’ — they are intrinsically inter-dependent. But, they are also salient and useful. One practical solution, then, is to accept those inter-dependencies when modelling, and to seek an optimal combination of those inter-dependent attributes for the forehead, or the sides of the nose, and other places where those inter-dependencies are found. Given that a measure of the ‘goodness of fit’ between the model mesh and a data mesh can be computed (giving an indication of how well a given combination of values fits the data), it is conceivable, but beyond the scope of this study, that the process of setting attributes can be automated, removing an huge inefficiency and source of measurement error.

## 8.2 The Potential Genetic Basis for Facial Attributes

At the beginning phase of this work, when features and attributes were identified by facial region, there was a sense that they are ad hoc but salient. We see qualities such as the breadth of the nose at various points along its length, and the prominence of the cheekbones, or the slope of the forehead, and we note that those qualities vary with ethnicity. These specific facial qualities were isolated into discrete, local, and separate named attributes such as *SOR\_protrusion* and *SL\_thickness*.

The distinguishing of some parts of the nose or eyes as named features might just reflect a general tendency for people to isolate parts of a complicated object, name them, and describe the complexity by a sort of ‘divide and conquer’ approach. It could further be argued that the face is in fact a single, continuous form, and that any apparent distinc-

tions of one part of that surface are just in our minds. If these seemingly discrete facial features are governed by distinct genetic codes, that would provide a basis for their being regarded as separable and ‘real’ (Liu et al., 2012; Paternoster et al., 2012; Claes et al., 2014, but see Hallgrímsson et al., 2014).

A set of facial attributes for which candidate genes have been identified (Adhikari et al., 2016), and their close correspondence with those developed in this research (see Table 8.1). The study by Adhikari et al. (2016) was published after the initiation of this study, and it is exciting to see an independent suggestion that discrete attributes such as the width of the bridge of the nose (which is called *DSM\_width* here) may actually correspond to a discrete trait that is programmed genetically, and moreover, that different genes contribute towards different traits, such as *DSM\_width* and *ALA\_width*.

While faces have an enormous range of variation, that face space can correspond to a fixed, common set of attributes or traits. One hundred or so such dimensions, each with

*Table 8.1. Correspondence between facial attributes implicated in a genetic study (Adhikari et al., 2016) and the attributes of the Ethnicity Modeller.*

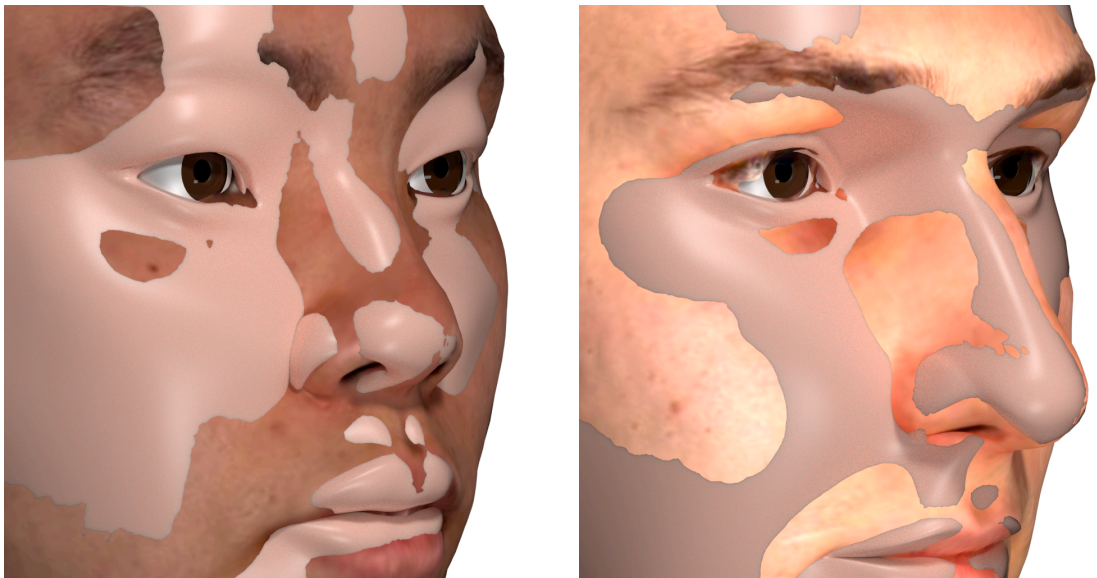
Genetic Trait	Region	Ethnicity Modeller Attribute
‘Columella inclination’	Nose	<i>TIP_inclination</i>
‘Columella protrusion’	Nose	<i>COL_drop</i>
‘Nasal bridge breadth’	Nose	<i>DSM_width</i>
‘Nasal root breadth’	Nose	<i>RAD_width</i>
‘Nose profile’	Nose	<i>DSM protrusion, RAD_protrusion, TIP_protrusion</i>
‘Nose protrusion’	Nose	<i>TIP_protrusion</i>
‘Nose tip shape’	Nose	<i>TIP_width; TIP_pointed</i>
‘Nose wing breadth’	Nose	<i>ALA_width</i>
‘Brow ridge protrusion’	Cranium	<i>SOR_protrusion</i>
‘Forehead profile’	Cranium	<i>FOR_slope</i>
‘Chin protrusion’	Jaw	<i>GNA_z</i>
‘Chin shape’	Jaw	<i>GNA_x</i>
‘Cheekbone protrusion’	Face	<i>CHK_protrusion</i>
‘Lower lip thickness’	Mouth	<i>IL_thickness</i>
‘Upper lip thickness’	Mouth	<i>SL_thickness</i>

a few distinguishable values, could account for the multitude of facial variations. Given this possibility, the set of facial traits that are genetically encoded might well correspond to those we see as salient.

### 8.3 An Economical Facial Representation Scheme

The end result of the modelling process is the creation of a file in JSON format (Section 6.5.1). This very modest file is basically a container for a dictionary of 77 entries, the important parts being the attribute-value pairs. In Figure 8.1 two individual faces are shown (an EAS male and an EUR male) as smooth surfaces accompanied with overlays of the stereo-photogrammetric data to show how well those models match the data on which they were based.

All models within the *EM* are the result of deforming a model mesh by the combined linear-weighted summation of 77 pairs of blendshapes (one pair for the two extrema for each attribute). Each blendshape is a low-polygon count mesh of merely 865 vertices



*Figure 8.1. Two individuals that were modelled within the EM were then stored as very small files (each a dictionary of 77 attribute-value pairs). Since all models share the same mesh geometry and the same framework of blendshapes, a model constitutes a very economical encoding of a given face.*

(the meshes are symmetrical so actually one could get by with half that number). By comparison, a stereo-photogrammetric data mesh consists of about 65,000 vertices plus the texture information.

In addition to providing an economical encoding scheme for faces, the attribute-value pairs are comparable across all models (for individuals or composites of individuals). In Table 8.2 the segment of the data pertaining to the 17 nose attributes are shown for the EAS and EUR individuals of Figure 8.1. The attributes were sorted by the absolute value of the difference in attribute value (in blue). The numerical differences between the two models quantified what are apparent by inspection, such as the EUR individual's nose protruding more than that of the EAS individual, as well as being narrower and longer, and so forth.

*Table 8.2. Comparison of the 17 attributes of the nose region for the two individuals modelled in Figure 8.1, sorted by descending absolute value of the difference in their values.*

Attribute	Description	Region	EAS	EUR	Difference
RAD_protrusion	radix protrusion (Z)	Nose	-0.75	0.78	1.53
DSM_protrusion	dorsum protrusion (Z)	Nose	-0.51	0.35	0.86
ALA_width	alar width al-al (X)	Nose	0.10	-0.71	0.81
DSM_width	dorsum width (X)	Nose	1.00	0.26	0.74
DSM_length	dorsum length n-sn (Y)	Nose	-0.15	0.48	0.63
TIP_protrusion	tip protrusion (Z)	Nose	-0.39	0.24	0.63
TIP_inclination	nasal tip inclination	Nose	0.87	0.26	0.61
SID_slope	sidewall slope (X)	Nose	-0.21	-0.75	0.54
TIP_drop	infratip lobule drop (Y)	Nose	0.15	0.61	0.46
COL_width	columella width (X)	Nose	-0.15	-0.51	0.36
TIP_pointed	tip pointed	Nose	-0.53	-0.33	0.20
COL_show	columellar show (Y)	Nose	0.54	0.74	0.20
TIP_width	tip width (X)	Nose	0.80	1.00	0.20
COL_drop	columella drop (Y)	Nose	0.16	0.29	0.13
ALA_drop	alar drop (Y)	Nose	0.07	0.18	0.11
ALA_contour	alar contour (sulcus)	Nose	0.51	0.53	0.02
RAD_width	radix width (X)	Nose	1.00	1.00	0.00

While the facial differences are revealed by Table 8.4, the ability to *visualise* these differences by animated interpolation is far more effective, especially in showing the simultaneous changes in these differences between the two types, which is better seen than read.

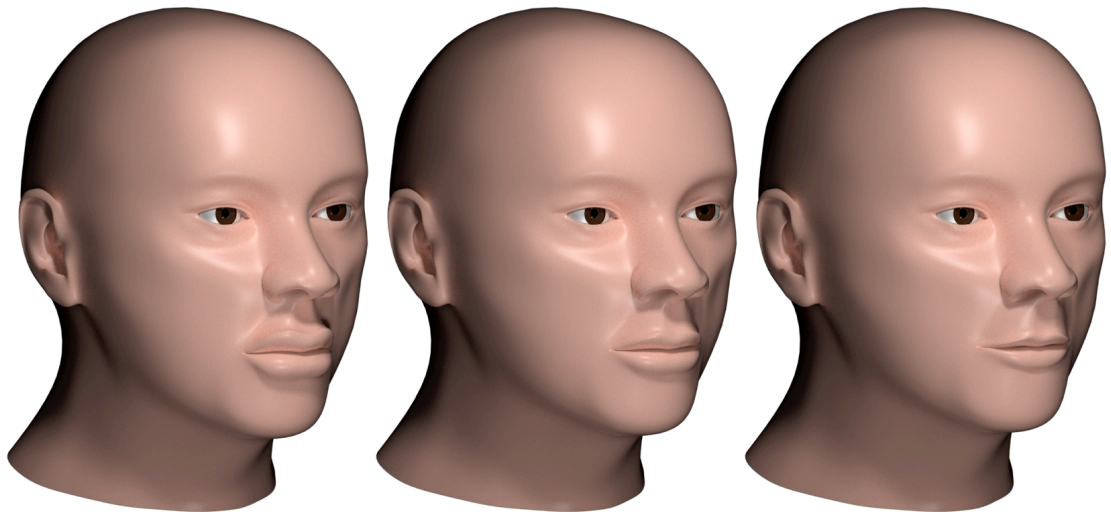
As is apparent in Figure 8.1 and Table 8.2, some attributes differ substantially in the two individuals (top rows of Table 8.2), while other attributes are negligibly different (bottom rows of Table 8.2). In a study of facial variation (in either a homogeneous population or across ethnotypes), the distribution of attribute-value pairs across a set of models could be subject to a principal components analysis to learn which attributes account for the greatest variation. PCA could be applied to those attributes within a given facial region, or to subsets of the attributes across regions, or to the entire set of facial attributes. This would provide insight into which attributes co-vary within each facial region (for example, does *SL\_thickness* co-vary with *IL\_thickness* in general, or only in some ethnicities?). PCA would be expected to reveal the expected patterns already known from anthropometric studies (e.g., that AFR faces have broad *ALA\_width* and broad *SL\_thickness*, and so forth). Perhaps unexpected correlations would be revealed as well, since the *EM* incorporates many attributes that are not conventionally analysed across ethnotypes.

#### 8.4 Experimental Investigation of Ethnicity Recognition

Referring back to Figure 7.5, the upper blend sequence of five frames shows interpolation between a generic AFR and an average of EUR, and the lower sequence between that AFR and an average of EAS. The three models AFR, EUR\_M20 and EAS\_M20 are each immediately identifiable as African, East Asian, and European, respectively. Now consider the blend sequence AFR-EUR, specifically. At the midpoint, the model is a 0.5 interpolation between the two. But what is the perceived ethnicity? In a forced-choice experiment (with effectively continuous gradations in blending between 0.0 and 1.0) does the point at which the face changes from appearing predominantly AFR (if forced to choose) to predominantly EUR occur at 0.5? It would be fascinating to use the *EM* to

explore what features are especially responsible for the recognition of one ethnotype relative to another.

In a related series, the interpolation provided by the *EM* can be used to create an ambiguous model such as a 0.5 interpolation of AFR and EUR (see Figure 8.2). Then by substituting an EUR mouth into the otherwise 50/50 blend (Figure 8.2, right), can that feature tilt the perception of the face as a whole as EUR? Likewise can an AFR mouth (Figure 8.2, left) shift the interpretation to AFR?



*Figure 8.2. The middle image shows a 0.5 interpolation between a generic AFR and an EUR average (EUR\_M20). (The specific AFR is a different individual than in Figure 7.5.). The models on the left and right have the same 50% AFR-EUR blend except on the left the mouth is 100% AFR and on the right the mouth is 100% EUR.*

Clearly the *Ethnicity Modeller* could be adapted to provide stimuli for a large number of perceptual studies of this sort.

## 8.5 Critical Evaluation

The goal of creating an EFS where the dimensions of that space correspond to intuitive and meaningful facial attributes, and those attributes are independent of one another, was only partly achieved. While many attributes are indeed orthogonal, others are *inter-de-*

pendent because their features share a common underlying structure. Some interdependence (as discussed in Section 6.4.4) is probably necessary if the nose is to be regarded as a ‘module’ that rides upon the maxillary bone. The shape of the nose is not entirely independent of the degree of protrusion of the mid and lower face due to the maxillary bone; for example, *TIP\_protrusion* is dependent upon *MAX\_protrusion*. This corresponds to how faces are constructed, but it does not allow these attributes to be orthogonal. Inter-dependence also arises when facial features spatially overlap. This problem was not adequately solved in this study, and would remain for future work. One solution to the coupling between attributes could be solved by re-factoring the set of attributes. For example, in the forehead (see Figure 6.18) there is clearly an interdependence between *FOR\_slope*, *FOR\_protrusion*, and *FOR\_curvature*. These are not, in retrospect, independent concepts, and therefore one should expect them to be create blendshape interference. Finding an alternative set of attributes that are more successfully orthogonal for the forehead, might be nontrivial. In fact, acceptance of some inter-dependency among attributes may be both practically and psychologically valid (we recognise that forehead slope depends somewhat on the degree of forehead protrusion and vice versa, and yet both are valid aspects of our description of a forehead).

The *Ethnicity Modeller* depends upon the user’s visual judgement and consequently, measurement error. Not only is the surface matching process laborious, it was found to limit the precision of the resulting model (Section 7.1.1).

A third weakness of the schema is the ad hoc process by which attributes are represented in terms of pairs of three-dimensional basis shapes. An inordinate effort was spent in adjusting the 77 pairs of meshes to minimise blendshape interference without resorting to *any* corrective blendshapes (a practical workaround in the animation industry but not applicable here; see Section 6.3.1). While a workable compromise was achieved, the process does not necessarily converge on a wholly-satisfactory set of blendshapes, nor would it be particularly repeatable by others. If the forehead region were to be refac-

tored, for instance, in order to be described by an alternative set of cranial attributes, one would be confronted with a fresh set of blendshape interference issues to be solved.

## 8.6 Summary

This study directly addressed the question of *how* to describe a face. Descriptions are usually provided in *written* form, and technical or specialised descriptions usually employ a specialised vocabulary for efficiency and precision. As discussed (Section 1.1), that vocabulary is often defined visually, i.e., with reference to illustrations or visual exemplars. A visual artist can communicate through depiction (e.g., abstractions, sketches, and pictographs) concepts that would be very difficult to evoke by any written description. These depictions evoke a concept by association, by our learned knowledge of what is being depicted, along with any written description. In this study, there is also a component of written description, and also a component of relying on examples. The written aspect is the local description of facial features by discrete *attribute-value* pairs, which are formatted into dictionaries in JSON format (Section 6.5.1).

For example, the supra-alar crease (or alar groove, termed the attribute *ALA\_contour* here), is defined anatomically primarily by its location (e.g., “crease located at cephalic border of ala”), while leaving implicit its characteristic shape (which follows the surface curvature of the ala as it merges into the dorsum). The shape and depth of this crease clearly varies with ethnicities as well as individually. This study captures a range for the attribute *ALA\_contour* in terms of an interpolation between two extremes. A specific nose may have the descriptor (*ALA\_contour* : 0.6), i.e., 60% between those extremes, where those extremes are represented three-dimensionally. Instead of words, the approach here relies on two concepts: examples (in this case of extremes of supra-alar crease) and interpolation within those extremes. This level of descriptive specificity is novel. As an artist, my challenge was to define such attributes as continuous-valued (as are all facial features), and to serve as independent descriptors of the overall face.



It is reasonable and natural for us to appeal to a visual example when defining some concept, particularly one that concerns shape. In this study, however, shape is not merely illustrated, but explicitly represented as a three-dimensional surface. So in fact one could illustrate the concept of *ALA\_contour* by animating that attribute within the *EM*, interpolating the blendshape between the two extreme basis shapes *ALA\_contour\_min* and *ALA\_contour\_max*. The definition of *ALA\_contour*, and the meaning of any specific value for this attribute, such as 0.6, is self-contained within this schema. One need not consult a medical dictionary or other supplemental sources to express this concept.

The practicality of implementing a purely orthogonal set of such facial attributes in terms of basis shapes was strained by the problem of blendshape interference. Some attributes appear to be intrinsically inter-related, but despite that non-orthogonality, they are natural and useful aspects of describing a face. Interestingly, genomic research may lend support to our intuition that faces are some encoded in terms of such discrete features.

Beyond the theoretical contribution of a schema for representing ethnic variation in the human face, the *implementation* of this scheme in terms of polygonal meshes opens the door to a large variety of practical three-dimensional applications, including forensic reconstruction, character design, ethnicity description, medical and anthropological education, and tools for face space visualisation and navigation. Moreover, in extending this approach beyond the domain of human faces and their variation, the concept of creating shape representations in terms of interpolation between extrema should have broader application, such as studying anthropological artefacts.

## 9. Conclusions and Future Directions

### 9.1 Contributions

#### 9.1.1 Regarding Ethnotypes and Face Spaces

Any representation scheme that can capture a range of faces would constitute a multidimensional ‘face space’, where the specific choice of dimensions for this face space depends on the specific application. The dimensionality of the space (however it is parameterised) is very great. A representation that supports efficient automatic recognition (such as an IFS) is not necessarily well-adapted to guiding a human observer. To explore facial variation across ethnicities, any proposed representational scheme needs to span the range of ethnic variation and to facilitate visualisation of similarities and distinctions between ethnicities.

*No ‘Average Face’.* The concept of an average face is fundamental to an IFS: it represents the IFS origin. A homogeneous population of individuals is presumed, with normally distributed facial variations relative to this average face. An individual face is then mapped into IFS according to its differences from the average along various dimensions of variation, so typical faces lie near the IFS origin, and similar faces project close to one other (Section 2.4.1). An IFS represents a space of different individual faces as distributed about the average face.

In contrast, the EFS in this study shares little with the conventional IFS. First, it is intended for description, not identification, and it makes no presumption about how facial variations are distributed (on the contrary, it’s purpose is to help investigate these variations). So there is no need to presume that the various dimensions on which ethnicities vary are normally distributed around some hypothetical average face for all of humankind, and the axes of an EFS need not map ethnotypes in terms of their differences relative to any such mean. The origin of the EFS does not represent the average value for each dimension; it is just the midpoint between the extremes of variation along that dimension.

*Defining an EFS Relative to the Extremes of Attribute Variation.* By defining each attribute in terms of the extremes for that attribute (instead of variation relative to presumed population mean), there is no requirement that a population show a central tendency or mean for that attribute; one could find any distribution of values for that attribute. More generally, if these attributes are regarded as EFS dimensions, there is no special significance to the origin. While an IFS is unbounded, and EFS is bounded.

Cluster analysis of the human genome supports adopting a geographical scheme for naming ethnotypes. This study focussed upon the three broad types African (AFR), East Asian (EAS) and European (EUR). Each of these ethnotypes could be further distinguished into sub-ethnotypes by narrowing the geographical region. An ethnotype defined at any such geographical scale exhibits large variations in the faces of that population, which limit how precisely the ‘averaged’ face of any ethnotype can be measured or described. An EFS does not map an ethnotype to a precise point, therefore. Rather, individuals of a given ethnotype would project to nearby points in an EFS.

*The Partitioning of an EFS into Subspaces.* This study identified 77 facial attributes, but recognised this total did not adequately capture the facial variation in any facial region. If one regards each attribute as a EFS dimension (disregarding the stubborn fact that they are not in fact strictly orthogonal), then a 100 (or more) DOF space would hardly be navigable.

Efforts to reduce the EFS dimensionality through PCA, however, will lose the ‘semantic’ understandability of the individual dimensions. An alternative is consider only sub-spaces, i.e., explore facial variation in only one facial region at a time. Fortunately, the geometric complexity of the face is primarily contained within spatially-separate facial regions, with few facial characteristics defined that spanning across regions. Given this natural partitioning, an EFS can then be regarded as composed of independent subspaces for the nose, eyes, and so forth.

A far greater simplification of an EFS is provided by reducing the space from *all* possible configurations for a given region to only those representative of various ethnicotypes. For example, the nose region contributes a total of 17 attributes (Section 5.3), and while the space of possible nose shapes is theoretically astronomical (Figure 6.16), those attributes are highly correlated for a given ethnicotype. The *observed* range of nose shapes within a given ethnicotype is far more constrained (Figure 5.9) and could be represented by the sample mean of the 17 nose attribute values plus their standard deviations (see Table A.3.4). For purposes of ethnicity exploration, therefore, the *EM* provides the basis for visualising the summary statistics of different ethnicities by exemplars, and ethnic differences can be appreciated visually by interpolating between, and among, these exemplars (Appendix A1.3 and Figure A1.3).

### 9.1.2 Shapes and Basis Shapes

*Facial Shape Descriptors.* While few anthropometric (dimensional) measurements, or ratios of such measurements, reliably distinguish between ethnicotypes, ethnic facial variation can be represented successfully in terms a set of discrete *shape* features, each with associated continuous-valued attributes. The attributes are implemented by three-dimensional basis shapes that permit spatial additivity to approximate a range of facial types. The attribute-value pairs form a representational scheme for ethnicotypes.

*Linear-Weighted Summation of Shape Attributes.* In PCA, faces are synthesised by linear-weighted summation of a set of orthogonal basis functions. Linear-weighted summation also underlies the well-established technique of blendshape deformation in facial animation (Section 6.3.1), and was also adopted for this study. But unlike the orthogonal ‘eigenmeshes’ mathematically derived from PCA, the manually-sculpted blendshape targets representing facial expressions in digital character animation are seldom orthogonal. Corrective blendshapes are often required to counteract blendshape interference, often by adding more blendshapes that compose facial expressions (Section 6.3.1). The considerable manual effort required to produce a workable set of blendshapes is clearly a worthwhile cost to the animation industry: linear-weighted summation is a very convenient

and powerful mechanism for mesh deformation. It is also capable of very precise displacements of individual vertices, as necessary in this study to create composites of facial attributes.

In this study, the blendshape targets directly represent facial attributes, hence blendshape interference cannot be controlled by adding hidden corrective blends to the *Ethnicity Modeller*. Blendshape interference could only be minimised by the painstaking refinement of the pair of basis shapes representing each attribute. It was eventually worth the effort, as it led to a satisfactory demonstration of a novel degree of what could be termed ‘attribute additivity’.

*Model Fitting.* The *EM* demonstrated an ability to parametrically match the surface geometry of a range of faces (Section 6.5.3). A stereo-photogrammetric data mesh of either an individual face or a delineated and composited average of multiple individual faces was imported into the *EM* and aligned spatially with the model mesh at two points (the corneas), presuming that data mesh had been rotated such that the Frankfort plane was horizontal. The 77 attributes were then adjusted until the surfaces of the model mesh and the data mesh coincided to within 2 mm (i.e., less than one standard deviation of the population samples). The set of attribute values that created this fit then constitutes a parametric model. Those attributes corresponding to conventional anthropometric measurements were calibrated in millimetres and showed a close correspondence with published anthropometric measurements for EAS and EUR ethnotypes (Section 7.1.2).

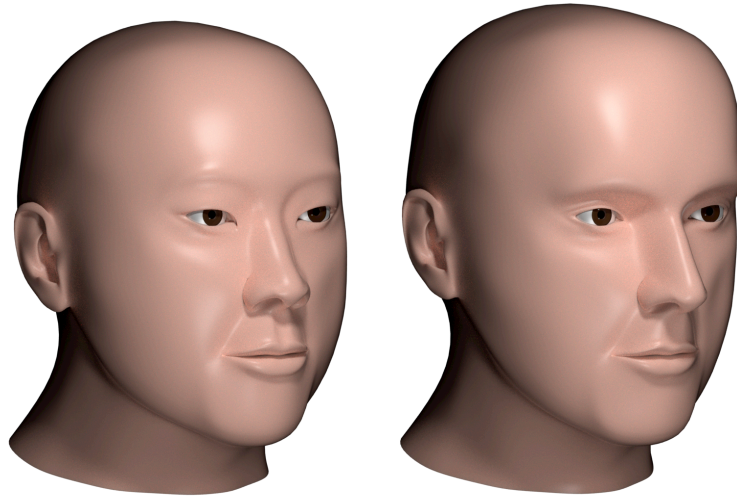
### 9.1.3 Visualising Ethnic Differences

*Visualising Facial Attributes.* Given a common set of basis shapes, a parametric model for a given ethnicity can be created and efficiently encoded in terms of a set of basis shape weights. Each basis shape provides a reference for the extreme of that attribute. Two models can then be compared in terms of their assigned weights for that attribute. Those shape attributes that correspond to conventional anthropometric landmarks can be calibrated and used for quantitative comparison across ethnicity models. A set of local

shape attributes is sufficient to model the averaged face of an ethnotype, yet not generally insufficient to capture the nuances of variation across individual faces. In fact, during development of the basis shapes, stereo-photogrammetric data of distinctive individuals were useful in ensuring that the basis shapes reliably represented extremes of various shape attributes.

*A Descriptive Representation of Facial Shape Attributes.* Complex objects are described by decomposing the whole into its parts, then describing the parts. For a manufactured object such as an automobile, the whole is literally made up of those parts. But for a biological form such as a face, the choice of what to call a ‘part’ is more a matter of convenience and convention. We tend to identify some facial features based on osteological or soft tissue landmarks. This study also followed the conventional strategy of first considering the face as composed of regions (eyes, nose, etc.) then within each region, to consider the features of that region, and then for each feature, to consider what properties are important to make explicit. While dimensional attributes (lengths, widths, etc, usually measured between anthropometric landmarks) are of course part of the description of a face, this study focussed on a set of shape attributes. While ad hoc and arbitrary, they are consistent with, and extend, traditional medical and anatomical terminology, and form the basis for the representation of the shape of a whole face from these parts.

*Visualisable and Quantifiable.* Local shape descriptors can be represented by basis shapes with sufficient specificity to reconstruct (i.e., match) facial geometry across a range of ethnotypes, to within the uncertainty that is imposed by individual variations. The shape attributes are local, comprehensible, and consistent with the traditional anatomical nomenclature and literature. The model (the set of attribute-value pairs) can then be used to deform a mesh for visualisation purposes, including interpolations between models, extrapolations of one model relative to another, and limited blends where selected subsets of all facial attributes are interpolated.



*Figure 9.1. The average of 20 EAS males (left) and that of 20 EUR males (right) differ significantly in many facial attributes. In side-by-side presentation reveals the major differences become apparent by shifting attention between the two. A tabulation of the attribute values for the two models (Table A3.4) reveals more subtle differences that are not so readily noticed by comparing static images. Perhaps the most effective means to appreciate their ethnic differences is through observing an animated interpolation between the two models ([Wisetchat, 2018](#)).*

*Animated Interpolation.* While static images are conventionally used to present a pair of faces A and B for side-by-side comparison. That task requires shifting attention back and forth between the two in search for where, and in what manner, the two faces differ. An alternative is to watch one face as it transforms from A to B (Section 2.3.3). Visual attention is drawn to where the shape is morphing, as that is where the two shapes differ. The primary advantage of animated interpolation may actually be in revealing relationships among attributes that change simultaneously. For example, comparing EAS and EUR, the protrusion of the supraorbital ridge, radix, and dorsum all increase while simultaneously the height of the supraorbital ridge decreases (Table A3.4). While this and other differences can be discerned in comparing the two static images in Figure 9.1, animated interpolation between the two is far more effective in revealing ethnic differences.

#### 9.1.4 Extending the Notion of Description by Example

The description and visualisation of ethnic variation in the human face bridges several disciplines (technology from computer graphics, nomenclature and measurements from medicine and anthropology, and visualisation and animation from graphic arts). While the word ‘description’ appears frequently in this study, it might seem to some readers that a description is not a *proper* description unless it is in the form of text. Instead, facial attributes were said to be ‘described’ by an attribute-value pair (a name and a number) and two associated three-dimensional basis shapes to define the range of that attribute. This notion of ‘description’ abandons the 18th century practice of creating a passive written discourse, and adopts the very current form of communication where graphical icons and symbols (pictographs such as emoji) are increasingly used to represent concepts, partly for their brevity, but more for their ability to convey some ideas more effectively than words. The use of three-dimensional shapes in the formation of descriptions would not have been feasible prior to the availability of the technology to create graphical visualisations.

## 9.2 Limitations

Some of earlier tasks of this study were sufficiently straightforward they built optimism about how much could eventually be achieved. An initial phase of photography resulted in face images for various ethnicities which were then delineated and averaged. Movie clips were then made that blend from one ethnicity to another, which readily revealed ethnic differences for each of the various facial regions (Chapter 5). Experimenting with alternative types of deformer in Maya and a review of animation industry practices soon led to choosing blendshape deformer for this study. A polygonal base mesh for the whole face was then sculpted, with edge loops and details added to support the various facial features. Copies of that base mesh were then sculpted to represent the extremes of each facial attribute. At this point, however, progress slowed as problems of blendshape interference began to appear, and the base mesh itself needed revision (causing revision of all basis shapes), the additional attributes needed to be added in response to attempting to adequately model different ethnicities. The *Ethnicity Modeller* was beginning to show



promise, but the initial goal of modelling a broad variety of ethnicities was replaced with learning how to effectively represent facial attributes by ‘quasi-orthogonal’ basis shapes.

It had been expected that, once the *EM* was implemented, different ethnotypes could be modelled using two-dimensional ‘image planes’ (e.g., an image-averaged ethnotype sample) as a template introduced into the Maya scene. But the significant problems inherent in attempting to match a three-dimensional mesh to a two-dimensional image (Section 6.5.2) meant the photograph catalog, while useful for identifying attributes (Chapter 5) was not used to model ethnotypes (Chapter 7). Instead, this study relied on more directly comparable, three-dimensional stereo-photogrammetric mesh data. While only two broad ethnotypes (EAS and EUR) were available as mesh data, the ability to model individuals of these ethnicities as well as averaged composites was very valuable (Appendix A3.4).

The study thus changed its emphasis at midpoint towards representational issues, and consideration of what constitutes a ‘description’ of a face beyond the familiar notions of measurements and specialised terminology. Where to go from here? Experience in creating the *Ethnicity Modeller* suggests various future directions of investigation.

### 9.3 Future Directions

In reflecting upon where to proceed from here, it is useful to start with what are clearly the most promising aspects of this schema. These include using the descriptors as a basis for encoding a face description, and the ease of extensibility of the user interface to create a broad variety of face description applications, as discussed next.

#### 9.3.1 Creating Derivative Applications

The *Ethnicity Modeller* (Chapter 6) was created to provide a concrete proof-of-concept demonstration (Section 7.1). The core architecture of the *EM* can repurposed to create derivative applications (see Sections 9.3.2-9.3.4). To see how this could be accom-

plished, first note that the *EM* consists of two separable components, one for visualisation and one for control.

The visualisation component is implemented as an Autodesk Maya scene graph containing the polygonal mesh `base` which is deformed by a ‘master’ blendshape node (`combined_blend`) which is in turn deformed by its connection to 77 additional blendshape nodes (one per attribute, of course, Section 6.3.1). The blendshape nodes are assigned interpolation values via the Python script `Model.py` (Appendix 2.3), which also reads and writes the JSON-format models. This entire scene graph (the polygonal mesh, two eyeballs, and associated material properties, cameras and lights) and scripting could readily be replicated in many other three-dimensional software systems such as Autodesk 3ds Max or Maxon Cinema 4D and Unity Technologies Unity. While Maya was chosen for development of the proof-of-concept implementation for this study, a derivative application would more appropriately be written in Unity or similar real-time engine.

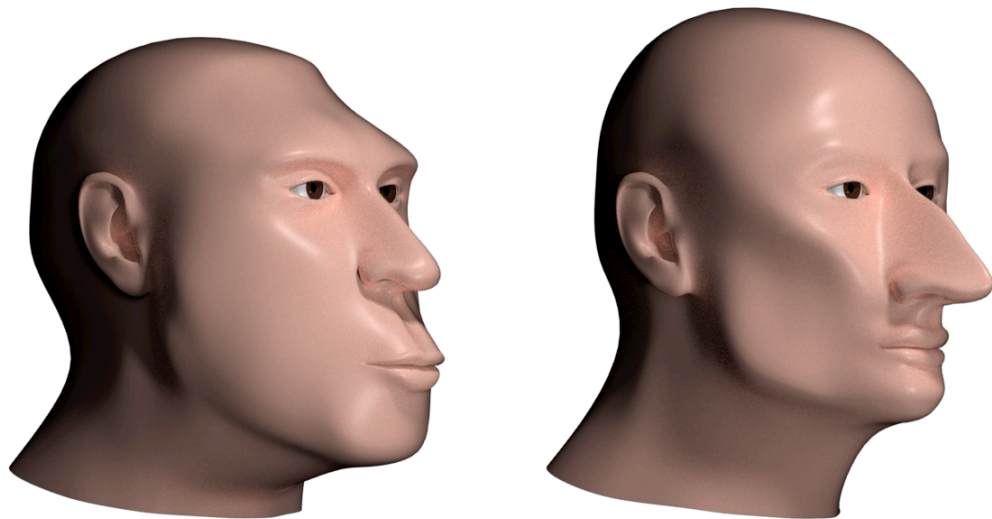
The control component consists of the user interface (`ModellerUI.py`, Appendix 2.4) and its associated configuration and launcher scripts (Appendices 2.1 and 2.2). The user interface was implemented in the widget SDK in `maya.cmds`, provided by Maya. The existing interface was intended solely to test modelling precision and accuracy, and to demonstrate navigation between different faces using linear interpolation.

If one were to create some derivative application, an entirely new UI would be tailored to that specific application, as well as modifying the visualisation component as needed for that application (for example, the face could readily be viewed in augmented reality). The code base in Appendix 2 reflects a fairly clean distinction between the visualisation and control components. The same format of data model (the JSON format file) could support many derivative applications. A given model would be read and its attribute values used to deform the face mesh, then if those values are edited interactively, that data could then be saved back as a model. Multiple such models could be read and interpolat-

ed. In this manner the same architecture can support many applications. Some prospective applications of the *Ethnicity Modeller* are sketched below.

### 9.3.2 Character and Caricature Design

Recall that Figure 7.1 showed some extreme combinations of facial attribute values. The intention was to search for blendshape interference. But of course those combinations of extreme attribute values result in grotesque faces. More moderate and more plausible combinations of facial features, can be created very readily through the *EM* user interface (see Figure 9.2). The manual process of character design could readily be automated through the introduction of heuristics for combining randomised attribute values within specified bounds in order to satisfy various intended qualities for the character. Fully-procedural face generation would have potential realtime applications where usually character design involves skilful use of three-dimensional sculpting tools such as McNeel Rhino, Pixologic ZBrush, Autodesk Mudbox, or Singular Inversions FaceGen.



*Figure 9.2. Two examples of caricatures created through the EM interface. The manual, interactive effort could be replaced by a purely procedural process for automated character creation.*

The above discussion concerns only the sculpting of the polygonal mesh. Current techniques for texture, bump, and geometry mapping make photo-realistic rendering of human skin possible. Moreover, particle systems permit the simulation of very natural-appearing hair, eyebrow, and eyelashes. These techniques could of course be applied to greatly enhance the realism of the face in the *Ethnicity Modeller*. Moreover, the addition of skin and hair properties (and eye colour, etc.) would permit controlling those ethnic variables as well. The polygonal mesh developed for this study already provides various sulci and skin folds (e.g., around the eyes) explicitly as polygonal geometry. It was be a straightforward matter of following standard industry practices to add fine skin details, and graduations in skin pigmentation and texture (such as apparent in Figures 4.1, 5.2, and 5.4). To be in keeping with this study, each such addition should be regarded a separate facial attribute.

### 9.3.3 Visualisation by Navigation

Since different faces project to different points of a face space, exploring different faces might be thought of as traveling through an Ethnicity Face Space (EFS). In a straightforward way the *EM* does provide a control panel for EFS navigation. But, rather than controlling a vehicle that moves in one dimension (a train), or two dimensions (a car), or three dimensions (an aircraft), the *EM* provides controls for moving in 77 dimensions. When adjusting one attribute at a time, the user of the *EM* travels along just one EFS dimension at a time, or, when interpolating between two model will travel along a straight-line path from one face to another, shifting proportionally along all 77 dimensions at once.

The idea of navigating through even three EFS dimensions at once is hard to imagine. Consider the effort that went into creating useful colour pickers, which navigate a merely three-dimensional space (e.g., HSV or RGB). It is sometimes surprisingly difficult to navigate a conventional colour space, for instance to interpolate between one colour and another by traveling along *three* dimensions simultaneously. Instead, it is commonplace to just pick a colour from a simplified two-dimensional palette of colour samples. Like-

wise, few ideas other than the equivalent of colour picking have been proposed for how to display alternative faces (see Section 2.4.3). One novel navigation tool comes to mind however, due to the geographical distribution of ethnicities, discussed next.

The globe (our planet) can be navigated by varying two degrees of freedom (latitude and longitude), and one could imagine watching a human face representative of the ethnicity that under the ‘cursor’ change in appearance as one traveled across the globe. Since ethnicities vary geographically (Section 2.1.2), and nearby ethnicities vary in many facial dimensions while sharing many others, could navigating across the globe help one to appreciate this complex variation by navigating geographically (i.e., in only the two degrees of freedom of latitude and longitude)? Similarly, suppose one considers two geographical regions A and B, separated by some significant geographical distance, with corresponding ethnotypes  $E_A$  and  $E_B$ . To what extent do the populations living between A and B reflect interpolations between  $E_A$  and  $E_B$ ? For instance, do the populations that live half way between these two populations represent 50/50 blends of the features of those two ethnotypes? It would not be expected to hold exactly, of course, but traversing the globe to investigate ethnic differences might reduce navigating the very high dimensional space of an EFS to simply navigating the two globe dimensions of latitude and longitude.

Applications of EFS navigation could be found in education, such as interactive displays for museums and educational software apps for mobile devices. Concepts such as ethnic diversity and human evolution can probably be explored more effectively through active interactive navigation than through passive means. Tools for creating a face out of ethnic ‘parts’ can help appreciate how the appearance of the overall face is more than the sum of its parts.

Upon reflection it is surprising that we are all aware of how the human face reflects its common origins as well as our diversity. Despite their individual differences, the individuals within each geographic locale share facial characteristics indicative of their eth-

nicity. The ethnic variation of the human face is seldom explicitly addressed in education. Faces tell about origins and cultures, and the language with which a face tells this story should be taught. What would constitute a descriptive language to express and communicate these differences? It would thus be of value to foster the appreciation of the face as telling the story of the commonality of all of humankind and the diversity in our global distribution (Wisetchat et al., 2018).

From the current study I would suggest a language not of words but of shapes, specifically three-dimensional shapes. Modern technology enables immersive visualisation of three-dimensional shape in compelling ways that facilitate our learning a language with which to describe faces. A pedagogical application could help develop an appreciation towards facial attributes, then the concept of a face space of possible combinations of attributes. That space could then be explored interactively to appreciate different specific ethnotypes, and finally, to develop a greater appreciation for the characteristics of one's own face in the context of global diversity.

#### 9.3.4 Other Domains

A final illustration of an application based on the *Ethnicity Modeller* is one of particular interest to me. One of the motivations for this study was my appreciation of the art history, cultural history, and aesthetics of the Buddha statues as they evolved from India and spread through South and Southeast Asia (Section 2.1). A lexicon of shape features specific to the Buddha statue had been developed by art historians and physical anthropologists, and used to infer the evolution of the statues (Stratton and Scott, 2004; Marwick, 2013).

A visualisation of the complex shape changes between pairs of Buddha styles was provided through digital animation (Wisetchat, 2013b). In that study the modelling of each of several different Buddha styles was achieved by 'manually' sculpting copies of a common polygonal mesh to resemble idealisations of the different styles. Interpolation from one model to another was achieved by straightforward blendshape interpolation of

the entire mesh; there was attempt to create models of different Buddha styles by different combinations of features. It would be informative to re-model the various sculptural styles of Buddha statues as different combinations of separate, discrete features whose attributes are implemented by basis shapes, and to re-approach the evolution of their complex stylistic difference in terms of linear combinations of their attributes.

This idea generalises: a family of complex objects that share a common structure (such as a similar arrangement of parts, shapes, or features) while differing in the specific properties or attributes of those features, could be modelled as different linear combinations of the attributes. Each attribute would be represented by a pair of basis shapes to cover the expected range of variation of that property for that feature. Then, rather than just attempt to model (i.e., replicate or depict) a given object, it could be both modelled and described by this vocabulary, and its differences from other objects in this family visualised through animated interpolation.

## References

- Adhikari, K., Fuentes, M., Quinto-Sanchez, M., Acuna-Alonzo, V., Jaramillo, C., Arias, W., Pizarro, M., Barquera, Lozano R., Macin Perez, G., Gomez-Valdes, J., et al. (2016). A genome-wide association scan implicates DCHS2, RUNX2, GLI3, PAX1 and EDAR in human facial variation. *Nature Communications*, 7:10815. pmid:26926045.
- Allanson, J.E. (1997). Objective techniques for craniofacial assessment: what are the choices? *American Journal of Medical Genetics*, 70(1), 1-5.
- Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(2), 155-194.
- Atiyeh, B.S. and Hayek, S.N. (2008). Numeric expression of aesthetics and beauty. *Aesthetic Plastic Surgery* (2008) 32:209–216 DOI 10.1007/s00266-007-9074-x.
- Bamshad, M., Wooding, S., Salisbury, B.A., and Stephens, J.C. (2004). Deconstructing the relationship between genetics and race. *Nature Reviews Genetics*, 5(8), 598-609.
- Bashour, M. (2006). History and current concepts in the analysis of facial attractiveness. *Plastic and Reconstructive Surgery*, 118(3), 741-756.
- Bekerman, I., Gottlieb, P., and Vaaiman, M. 2014. Variations in eyeball diameters of the healthy adults. *Journal of Ophthalmology*, 2014, Article ID 503645.
- Benson, P.J., and Perrett, D.I. (1991). Synthesizing continuous-tone caricatures. *Image & Vision Computing*, 9, 123-129.
- Benson, P.J. and Perrett, D.I. (1993). Extracting prototypical facial images from exemplars. *Perception*, 22, 257–262.



- Bergeron, P. and Lachapelle, P. (1985). Controlling facial expression and body movements in the computer-generated short 'Tony de Peltrie', *Siggraph Tutorial Notes*, New York: ACM Press.
- Bernier, F. (2001). A new division of the Earth. *History Workshop Journal*, 51, 247-250.
- Bibliowicz, J. (2005). An automated rigging system for facial animation. Master's Thesis, Cornell University.
- Blanz, V. and Vetter, T. (1999). A morphable model for the synthesis of 3D faces. In: *Proceedings of the 26th annual conference on Computer graphics and interactive techniques* (pp. 187-194). ACM Press/Addison-Wesley Publishing Co.
- Blanz, V. and Vetter, T. (2003). Face recognition based on fitting a 3d morphable model. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(9), 1063-1074.
- Blanz, V., Scherbaum, K., & Seidel, H.P. (2007). Fitting a morphable model to 3D scans of faces. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on Computer Vision* (pp. 1-8). IEEE.
- Blumenbach, J.F. (1795). *Elements of physiology*. Johann Friedrich; Caldwell, Charles (Trans). Vols. 1 and 2. Unknown Publisher 476 pp.
- Boehm, B.W. (1988). A spiral model of software development and enhancement. *Computer*, 21(5), 61-72.
- Bookstein, F.L. (1991). *Morphometric tools for landmark data: Geometry and biology*. New York: Cambridge University Press.

- Brothwell, D.R. and Harvey, R.G. (1965). Facial variation, with special reference to the people of Tristan Da Cunha. *The Eugenics Review*, 57(4), 167.
- Burt, D.M. and Perrett, D.I. (1995). Perception of age in adult Caucasian male faces: computer graphic manipulation of shape and color information. *Proceedings of the Royal Society Series B*, 259(1355), 137-143.
- Burton, A.M., Bruce, V., and Dench, N. (1994). What's distinctive about a distinctive face. *Quarterly Journal of Experimental Psychology*, 47A, 119-141.
- Burton, A.M., Jenkins, R., Hancock, P.J., and White, D. (2005). Robust representations for face recognition: The power of averages. *Cognitive Psychology*, 51(3), 256–284.
- Burton, A.M. and Vokey, J. (1998). The face space typicality paradox: Understanding the face-space metaphor. *Quarterly Journal of Experimental Psychology*, 51A, 475-483.
- Busey, T.A. (1998). Physical and Psychological Representations of Faces: Evidence from Morphing. *Psychological Science*, 9(6), 476-483.
- Cavalli-Sforza, L.L., Menozzi, P., and Piazza, A. (1994). *The history and geography of human genes*. Princeton: Princeton University Press.
- Chen, T.P.G. (2003). 1,001,001 faces: A configural face navigation interface (Doctoral dissertation, University of British Columbia).
- Chen, T.P.G. and Fels, S. (2004). Exploring gradient-based face navigation interfaces. In *Proceedings of Graphics Interface 2004* (pp. 65-72). Canadian Human-Computer Communications Society, 2003.

- Chandra, H.J., Ravi, M.S., Sharma, S.M., and Prasad, B.R. (2012). Standards of Facial Esthetics: An Anthropometric Study. *Journal of Maxillofacial and Oral Surgery*, 11(4), 384-389.
- Choe, K.S., Sclafani, A.P., Litner, J.A., Yu, G.P., and Romo, T. (2004). The Korean American woman's face: anthropometric measurements and quantitative analysis of facial aesthetics. *Archives of Facial Plastic Surgery*, 6(4), 244-252.
- Choe, K.S., Yalamanchili, H.R., Litner, J.A., Sclafani, A.P., and Quatela, V.C. (2006). The Korean American Woman's Nose. *Archives of Facial Plastic Surgery*, 8(5), 319-323.
- Coetzee, V., Grief, J.M., Stephen, I.D., and Perrett, D.I. (2014). Cross-cultural agreement in facial attractiveness preferences: The role of ethnicity and gender. *PLoS ONE*, 9(7): e99629. doi:10.1371/journal.pone.0099629.
- Claes, P., Liberton, D.K., Daniels, K., Rosana, K.M., Quillen, E.E., Pearson, L.N., et al. (2014). Modeling 3D Facial Shape from DNA. *PLoS Genetics*, 10(3): e1004224. doi: 10.1371/journal.pgen.1004224.
- Darwin, C. (1859). *The origin of species by means of natural selection or the preservation of favoured races in the struggle for life*. J.W. Burrow, ed. London: Penguin Books.
- Dawei, W., Gouging, Q., Mingle, Z., and Farkas, L.G. (1997). Differences in horizontal, neoclassical facial canons in Chinese (Han) and North American Caucasian populations. *Aesthetic Plastic Surgery*, 21(4), 265–269.
- DeBruine, L.M. and Tiddeman, B.P. (2017). WebMorph. Available at: <https://webmorph.org>. [Accessed July 7, 2018].

- Diskul, S. (1991). *Art in Thailand: A Brief History*. Bangkok: Amarin Printing Group.
- Douglas, T.S., Menthes, E.M., Vaughan, C.L., and Viljoen, D.L. (2003). Role of depth in eye distance measurements: comparison of single and stereo-photogrammetry. *American Journal of Human Biology*, 15, 573-578.
- Dawei, W., Guozheng, Q., Mingli, Z., and Farkas, L.G. (1997). Differences in horizontal, neoclassical facial canons in Chinese (Han) and North American Caucasian populations. *Aesthetic Plastic Surgery*, 21(4), 265-269.
- Drake, A.G. and Klingenberg, C.P. (2008). The pace of morphological change: historical transformation of skull shape in St. Bernard dogs. *Proceedings of the Royal Society of London. B Biological Science*, 275, 71–76.
- Dryden, I.L., and Mardia, K.V. (1998). *Statistical shape analysis, with applications in R (2nd addition)*. Chichester: John Wiley & Sons.
- Duckworth, W.L.H. (1904). *Morphology and anthropology: a handbook for students*. Cambridge: University Press.
- Dürer, A. (1981). *The art of measurement*. San Francisco: Alan Wofsy Fine Arts.
- Elewa, A.M.T. (2010). *Morphometrics for nonmorphometricians (Vol. 124)*. Berlin: Springer.
- Elmqvist, N., Dragicevic, P., and Fekete, J.D. (2008). Rolling the dice: Multidimensional visual exploration using scatterplot matrix navigation. *IEEE Transactions on Visualization and Computer Graphics*, 14(6), 1539-1148.

- Enciso, R., Shaw, A.M., Neumann, U., and Mah, J. (2003). Three-dimensional head anthropometric analysis. In: *Medical Imaging 2003* (pp. 590-597). International Society for Optics and Photonics.
- Ersotelos, N., and Dong, F. (2008). Building highly realistic facial modeling and animation: a survey. *The Visual Computer*, 24(1), 13-30.
- Fang, F., Clapham, P.J., and Chung, K.C. (2011). A systematic review of interethnic variability in facial dimensions. *Plastic and Reconstructive Surgery*, 127(2), 874–881.
- Farkas, L.G. (1981). *Anthropometry of the head and face in medicine*. New York: Elsevier.
- Farkas, L.G. (1994). Anthropometry of the head and face in clinical practice. In: L.G. Farkas, ed. *Anthropometry of the Head and Face 2nd edition*, New York: Raven Press, pp. 71-77.
- Farkas, L.G., and Cheung, G. (1981). Facial asymmetry in healthy North American Caucasians: an anthropometrical study. *The Angle Orthodontist*, 51(1), 70-77.
- Farkas, L.G., Hreczko, T.A., Kolar, J.C., and Munro, I.R. (1985). Vertical and horizontal proportions of the face in young adult North American Caucasians: revision of neoclassical canons. *Plastic and Reconstructive Surgery*, 75(3), 328-337.
- Farkas, L.G., Forrest, C.R., and Litsas, L. (2000). Revision of neoclassical facial canons in young adult Afro-Americans. *Aesthetic Plastic Surgery*, 24(3), 179-184.
- Farkas, L.G., Katic, M.J., and Forrest, C.R. (2005). International anthropometric study of facial morphology in various ethnic groups/races. *Journal of Craniofacial Surgery*, 16(4), 615-646.

- Fisher, R. E. (1993). *Buddhist Art and Architecture*. New York: Thames and Hudson, Ltd.
- Franco, F.C., Arroyo, F.M., Vogel, C.J., and Quintão, C.C. (2013). Brachycephalic, dolichocephalic and mesocephalic: is it appropriate to describe the face using skull patterns? *Dental Press Journal of Orthodontics*, 18(3), 159-163.
- Galton, F. (1878). Composite portraits. *Journal of the Anthropological Institute of Great Britain and Ireland*, 8, 132–144.
- Ghoddousi, H., Edler, R., Haers, P., Wertheim, D., and Greenhill, D. (2007). Comparison of three methods of facial measurement. *International Journal of Oral and Maxillofacial Surgery*, 36(3), 250-258.
- Gray, H. (1918). *Anatomy of the human body*. Philadelphia: Lea and Febiger.
- Hallgrimsson, B., Mio, W., Marcucio, R.S., Spritz, R. (2014). Let's Face It— Complex Traits Are Just Not That Simple. *PLoS Genetics*, 10(11): e1004724. doi:10. 1371/ journal.pgen.1004724.
- Hancock, P.J., Burton, A.M., and Bruce, V. (1996). Face processing: Human perception and principal components analysis. *Memory and Cognition*, 24(1), 26-40.
- Hanihara, T. (2000). Frontal and facial flatness of major human populations. *American Journal of Physical Anthropology*, 111(1), 105-134.
- Helmholtz, H.V. (1867, 1962). *Helmholtz's treatise on physiological optics*. J.P.C. Southall, ed. New York: Dover.

- Hertzum, M. and Jacobsen, N.E. (2001). The Evaluator Effect: A Chilling Fact About Usability Evaluation Methods. *International Journal of Human-Computer Interactions*, 13(4), 421–443.
- Holub, A., Liu, Y.H., and Perona, P. (2007). On constructing facial similarity maps. In *IEEE Computer Vision and Pattern Recognition, CVPR'07*, 1-8.
- Holzleitner, I.J., Hunter, D.W., Tiddeman, B.P., Seck, S., Re, D.E., and Perrett, D.I. (2014). Men's facial masculinity. When (body) size matters. *Perception*, 43, 1191-1202.
- Huxley, T.H. (1870). On the geographical distribution of the chief modifications of mankind. *The Journal of the Ethnological Society of London (1869-1870)*, 2(4), 404-412.
- Jalal, G., Maudet, N., and Mackay, W.E. (2015). Color portraits: From color picking to interacting with color. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 4207-4216). ACM.
- Jayaratne, Y.S., Deutsch, C.K., and Zwahlen, R.A. (2013). A 3D anthropometric analysis of the orolabial region in Chinese young adults. *British Journal of Oral and Maxillofacial Surgery*, 51(8), 908-912.
- Jeffries, J.M., Dibernardo, B., and Raunchier, G.E. (1995). Computer analysis of the African- American face. *Annals of Plastic Surgery*, 34, 318-322.
- Khambay, B., Nairn, N., Bell, A., Miller, J., Bowman, A., and Ayoub, A.F. (2008). Validation and reproducibility of a high-resolution three-dimensional facial imagine system. *British Journal of Oral and Maxillofacial Surgery*, 46(1), 27-32.

- Kendall, D.G. (1977). The diffusion of shape. *Advances in Applied Probability*, 9, 428-430.
- Keim, D.A. (2002). Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1), 1–8.
- Klingenberg, C.P. (2013). Visualizations in geometric morphometrics: how to read and how to make graphs showing shape changes. *Hystrix, the Italian Journal of Mammalogy*, 24(1), 15-24.
- Lam, V.B., Czyz, C.N., and Wulc, A.E. (2013). The brow-eyelid continuum: an anatomic perspective. *Clinics in Plastic Surgery*, 40(1), 1–19.
- Langlois, J.H. and Roggman, L.A., (1990). Attractive faces are only average. *Psychological Science*, 1(2), 115-121.
- Le, T.T., Farkas, L.G., Ngim, R.C., Levin, L.S., and Forrest, C.R. (2002). Proportionality in Asian and North American Caucasian faces using neoclassical facial canons as criteria. *Aesthetic Plastic Surgery*, 26(1), 64-69.
- Lewis, C. (1982). Using the “thinking-aloud” method in cognitive interface design (IBM Research Rep. No. RC 9265 [#40713]). Yorktown Heights, NY: IBM Thomas J. Watson Research Center.
- Lewis, J.R. (1994). Sample sizes for usability studies: Additional considerations. *Human Factors*, 36, 368–378.
- Lewis, J.P., Mooser, J., Deng, Z., and Neumann, U. (2005). Reducing blendshape interference by selected motion attenuation. In: *Proceedings of the 2005 Symposium on Interactive 3D Graphics and Games*, 25-29, New York: ACM.



- Lewis, J.P. and Anjyo, K.I. (2010). Direct-manipulation blendshapes. *IEEE Computer Graphics and Applications*, 30(4), 42-50.
- Lewis, M.B. (2004). Face-space-R: towards a unified account of face recognition. *Visual Cognition*, 11(1), 29-69.
- Lewontin, R.C. (1972). The apportionment of human diversity. *Evolutionary Biology*, 6, 381-398.
- Liu, F., Van Der Lijn, F., Schurmann, C., Zhu, G., Chakravarty, M.M., Hysi, P.G., Wollstein, A., Lao, O., De Bruijne, M., Ikram, M.A. and Van Der Lugt, A. (2012). A genome-wide association study identifies five loci influencing facial morphology in Europeans. *PLoS genetics*, 8(9), p.e1002932.
- Marwick, B. (2012). A cladistic evaluation of ancient Thai bronze Buddha images: six tests for a phylogenetic signal in the Griswold collection. In: M.L. Tjoa-Bonatz, A. Reinecke, and D. Bonatz, eds. *Connecting empires and states: selected papers from the 13th international conference of the European association of Southeast Asian archaeologists*. Singapore: National University of Singapore Press. pp. 159-176.
- Mayr, E. (1962). Origin of the Human Races. *Science*, 138: 420-422.
- McIntyre, A.H., Hancock, P.J., Kittler, J., and Langton, S.R. (2013). Improving discrimination and face matching with caricature. *Applied Cognitive Psychology*, 27(6), 725-734.
- Medlej, J. (2012). Drawing people. The artist's guide to anatomy, physical features, facial expressions, movement, ageing, and ethnotypes. Self-published. Available at <http://www.lulu.com/shop/joumana-medlej/drawing-people-ebook/ebook/product-20039566.html>.

- Metraux, A. (1950). United Nations Economic and Security Council statement by experts on problems of race, *American Anthropologist*, 53(1), 142-145.
- Ochs, M.W. and Buckley, M.J. (1993). Anatomy of the orbit. *Oral and Maxillofacial Surgery Clinics of North America*, 5, 419-429.
- Ofodile, F.A., Bokhara, F.J., and Ellis, C. (1993). The black American nose. *Annals of Plastic Surgery*, 31, 209-219.
- O'Neill, J.P. and da Vinci, L. (1983). *Anatomical Drawings from the Royal Library Windsor Castle*. The Metropolitan Museum of Art.
- Orvalho, V., Bastos, P., Parke, F.I., Oliveira, B., and Alvarez, X. (2012). A Facial Rigging Survey. In: *Eurographics State of the Art Report*, 183-204.
- O'Toole, A.J., Wenger, M.J., and Townsend, J.T. (2001). Quantitative models of perceiving and remembering faces: Precedents and possibilities. In: M.J. Wenger & J.T. Townsend, eds. *Computational, geometric, and process perspectives on facial cognition: Contexts and challenges*, 1-38, New York: Psychology Press.
- Parent, R. (2012). *Computer animation. Algorithms and techniques*. Waltham, MA: Morgan Kaufmann.
- Parke, F.I. (1972). Computer generated animation of faces. In: *Proceedings of the ACM annual conference-Volume 1* (pp. 451-457). ACM.
- Paternoster, L., Zhurov, A.I., Toma, A.M., Kemp, J.P., Pourcain, B.S., Timpson, N.J., McMahon, G., McArdle, W., Ring, S.M., Smith, G.D. and Richmond, S. (2012). Genome-wide association study of three-dimensional facial morphology identifies a

variant in PAX3 associated with nasion position. *The American Journal of Human Genetics*, 90(3), 478-485.

Pérez-Pérez, A., Alesan, A., and Roca, L. (1990). Measurement error: inter-and intra observer variability. An empiric study. *International Journal of Anthropology*, 5(2), 129-135.

Perrett, D.I., May, K.A., and Yoshikawa, S. (1994). Facial shape and judgements of female attractiveness. *Nature*, 368, 239-242.

Perrett, D.I., Lee, K.J., Penton-Voak, I.S., Rowland, D., Yoshikawa, S., Burt, D.M., Henzi, S.P., Castles, D.L., and Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature*, 394, 884–887.

Perrett, D.I., Burt, D.M., Penton-Voak, I.S., Lee, K.J., Rowland, D.A., and Edwards, R. (1999). Symmetry and human facial attractiveness. *Evolution and Human Behavior*, 20, 295-307.

Porter, J.P. (2004). The average African American male face. *Archives of Facial Plastic Surgery*, 6(2), 78-81.

Porter, J.P. and Olson, K.L. (2001). Anthropometric facial analysis of the African American woman. *Archives of Facial Plastic Surgery*, 3(3), 191-197.

Raitt, B. (2004). The making of Gollum. Presentation at U. Southern California Institute for Creative Technologies's *Frontiers of Facial Animation* Workshop, August.

Rennels, J.L., Bronstad, P.M., and Langlois, J.H. (2008). Are attractive men's faces masculine or feminine? The importance of type of facial stimuli. *Journal of Experimental Psychology. Human Perception and Performance*, 34(4), 884–93.

- Ritz-Timme, S., Gabriel, P., Obertová, Z., Boguslawski, M., Mayer, F., Drabik, A., Poppa, P., De Angelis, D., Ciaffi, R., Zanotti, B., and Gibelli, D. (2011). A new atlas for the evaluation of facial features: advantages, limits, and applicability. *International Journal of Legal Medicine*, 125(2), 301-306.
- Rosenberg, N.A., Pritchard, J.K., Weber, J.L., Cann, H.M., Kidd, K.K., Zhivotovsky, L.A., and Feldman, M.W. (2002). Genetic structure of human populations. *Science*, 298(5602), 2381-2385.
- Rowland, Jr., B. (1963). *The Evolution of the Buddha image*. Asia House Gallery Publication/Harry N. Abrams, Inc.
- Sarich, V., and Miele, F. (2003). *Race: The reality of human differences*. Boulder, CO: Westview Press.
- Scott, I.M.L., Pound, N., Stephen, I.D., Clark, A.P., and Penton-Voak, I.S. (2010). Does masculinity matter? The contribution of masculine face shape to male attractiveness in humans. *PLoS ONE*, 5(10): e13585. doi:10.1371/journal.pone.0013585.
- Sforza, C. and Ferrario, V.F. (2006). Soft-tissue facial anthropometry in three-dimensions: from anatomical landmarks to digital morphology in research, clinics and forensic anthropometry. *Journal of Anthropological Sciences*, 84, 97-124.
- Shepard, R.N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika*, 27(2), 125–140.
- Shepard, R.N. (1987). Towards a universal law of generalization for psychological science. *Science*, 237(4820), 1317-1323.

- Sim, R.S., Smith, J.D., and Chan, A.S. (2000). Comparison of the aesthetic facial proportions of southern Chinese and white women. *Archives of Facial Plastic Surgery*, 2(2), 113-120.
- Sherman D.D., Burkat, C.N., and Lemke, B.N. (1992). Orbital Anatomy and Its Clinical Applications. In: Tasman, W. and Jaeger, E.A. *Duane's Ophthalmology on CD-ROM, 2006 Edition*. Philadelphia: Lippincott Williams & Wilkins.
- Stephan, C.N., Penton-Voak, I.S., Perrett, D.I., Tiddeman, B.P., Clement, J.G., and Henneberg, M. (2005). Two-dimensional computer-generated average human face morphology and facial approximation. In: J.G. Clement and M.K. Marks. *Computer-graphic facial reconstruction*. New York: Elsevier, 105-127.
- Stephan, C.N. and Davidson, P.L. (2008). The Placement of the Human Eyeball and Canthi in Craniofacial Identification. *Journal of Forensic Sciences*, 53(3), 612-619.
- Steyvers, M. (1999). Morphing techniques for manipulating face images. *Behavior Research Methods, Instruments, and Computers*, 31(2), 59-369.
- Stratton, C. and Scott, M. (2004). *Buddhist Sculpture of Northern Thailand*. Chicago: Serindia Publications.
- Strauss, R.E. (2010). Discriminating groups of organisms. In: A.M.T. Elewa, ed. *Morphometrics for nonmorphometricians*. Berlin: Springer, pp. 73-91.
- Swennen, G.R. (2006). 3-D Cephalometric soft tissue landmarks. In: G.R. Swennen, et al., eds. *Three-dimensional cephalometry: a color atlas and manual*. Berlin: Springer Science and Business Media, pp. 186-226.

- Swennen, G.R., Schutyser, F.A., and Hausamen, J.E., eds. (2006). *Three-dimensional cephalometry: a color atlas and manual*. Berlin: Springer Science and Business Media.
- Taylor, J.R. (1997). *An Introduction to error analysis. The study of uncertainties in physical measurements*. Sausalito, CA: University Science Books. 327 pp.
- Thompson, D.A.W. (1915). Morphology and mathematics. *Transactions of the Royal Society of Edinburgh*, 50, 857–895.
- Thompson, D.A.W. (1917). *On Growth and Form*. Cambridge: Cambridge University Press.
- Tiddeman, B.P., Duffy, N, and Rabey, G. (2000). Construction and visualisation of three-dimensional facial statistics. *Computer Methods and Programs in Biomedicine* 63, 9-20.
- Tiddeman, B.P., Burt, D.M., and Perrett, D.I. (2001). Prototyping and transforming facial texture for perception research. *IEEE Computer Graphics Applications*, 21, 42–50.
- Tiddeman, B., Rabey, G., and Duffy, N. (1999). Synthesis and transformation of three-dimensional facial images. *IEEE Engineering in Medicine and Biology Magazine : the Quarterly Magazine of the Engineering in Medicine & Biology Society* 18(6), 64-9.
- Tracy, W. (1986). *Letters of credit. A view of type design*. London: Gordon Fraser Gallery, Ltd.
- Treves, A. (1997). On the perceptual structure of face space. *BioSystems*, 40(1-2), 189-196.

- Tuncel, U., Turan, A., and Kostakoğlu, N. (2013). Digital anthropometric shape analysis of 110 rhinoplasty patients in the Black Sea Region in Turkey. *Journal of Cranio-Maxillofacial Surgery*, 41(2), 98-102.
- Turk, M. and Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience* 3(1), 71-86.
- United Nations Statistics Division (2007). Composition of macro geographical (continental) regions, geographical sub-regions, and selected economic and other groupings, United Nations Statistics Division. Revised August 28, 2007. Available at: <https://unstats.un.org/unsd/methodology/m49/> [Accessed on line July 7, 2018].
- United Nations Department of Economic and Social Affairs and Population Division (2015). World Population Prospects: The 2015 Revision, Key Findings and Advance Tables. Working Paper, No. ESA/P/WP.241. New York, NY: United Nations Department of Economic and Social Affairs.
- Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology*, 43(2), 161-204.
- Valentine, T. (2001). Face-space models of face recognition. Computational, geometric, and process perspectives on facial cognition. In: M.J. Wenger & J.T. Townsend, eds. *Computational, geometric, and process perspectives on facial cognition: Contexts and challenges*, 83-113, New York: Psychology Press.
- Valentine, T., Lewis, M.B., and Hills, P.J. (2016). Face-space: A unifying concept in face recognition research. *The Quarterly Journal of Experimental Psychology*, 69(10), 1996-2019.

- Valenzano, D.R., Mennucci, A., Tartarelli, G., and Cellerino, A. (2005). Shape analysis of female facial attractiveness. *Vision Research*, 46(8-9), 1282-1291.
- Waters, K. (1987). A muscle model for animation three-dimensional facial expression. *ACM SIGGRAPH Computer Graphics*, 21(4), 17-24.
- Wiley, D.F., Amenta, N., Alcantara, D.A., Ghosh, D., Kil, Y.J., Delson, E., Harcourt-Smith, W., Rohlf, F.J., St. John, K., and Hamann, B. (2005). Evolutionary morphing. *Proceedings of the IEEE Visualization 2005 (VIS'05)*, 431–438.
- Winder, R.J., Darvann, T.A., McKnight, W., Magee, J.D.M., and Ramsay-Baggs, P. (2008). Technical validation of the Di3D stereophotogrammetry surface imaging system. *British Journal of Oral and Maxillofacial Surgery*, 46(1), 33-37.
- Wisetchat, S. (2011). Sukhothai: The Evolution of a distinctly Thai sculptural style. M.Phil. Thesis. The Glasgow School of Art, Digital Design Studio, University of Glasgow, Glasgow, UK.
- Wisetchat, S. (2013a). Visualizing the evolution of the Sukhothai Buddha. *Southeast Asian Studies*, 2(3), 559-582. Available at: <http://englishkyoto-seas.org/2014/01/vol-2-no3-sawitree-wisetchat/> [Accessed July 7, 2018].
- Wisetchat, S. (2013b). Digital modeling of Buddha sculptures. *Silpakorn University Journal of Social Sciences, Humanities, and Arts*, 13(2), 29-46. Available at: <https://tcithaijo.org/index.php/hasss/article/view/12559> [Accessed July 7, 2018].
- Wisetchat, S. (2018). Ethnicity Modeller. Available at: <http://sawitreedesigns.wixsite.com/index/exploring> [Accessed September 17, 2018].



- Wisetchat, S., DeBruine, L., and Livingstone, D. (2018). Exploring how faces reveal our ethnicity In: *iLRN 2018 Montana* (pp. 104-114). DOI: <https://doi.org/10.3217/978-3-85125-609-3-14>. Available at: <http://openlib.tugraz.at/download.php?id=5b35f09e4db1b&location=browse> [Accessed July 7, 2018].
- Witherspoon, D.J., Wooding, S., Rogers, A.R., Marchani, E.E., Watkins, W.S., Batzer, M.A. and Jorde, L.B. (2007). Genetic similarities within and between human populations. *Genetics*, 176(1), 351-359.
- Xing, S. Gibbom, V. Clarke, R. W. Liu. (2013). Geometric morphometric analyses of orbit shape in Asian, African, and European human populations. *Anthropological Science*, 121(1), 1-11.
- Zhang, Y. and Badler, N.I. (2006). Synthesis of 3D faces using region-based morphing under intuitive control. *Computer Animation and Virtual Worlds*, 17, 421–432.
- Zhang, Y. and Badler, N.I. (2007). Face modeling and editing with statistical local feature control models. *International Journal of Imaging Systems and Technology*, 17(6), 341–358.

## Appendix 1. User Interface

### A1.1 Initialising the Ethnicity Modeller

The Ethnicity Modeller is run within the Maya scene `modeler.ma` within the project `EthnicityModeller`. The transform node of the polygonal mesh representing the face is named `base`, and is in the `Render_Layer` display layer. This mesh is influenced by a blendshape deformer (`base_blendShape`) and a large set of blend targets, each corresponding to the implementation of an extreme of a facial attribute. These target meshes are organised by facial region and placed in corresponding display layers. They are protected in reference mode and normally invisible.

The Ethnicity Modeller is controlled via a user interface. Select the `EthnicityModeller` shelf and launch the user interface by clicking the `EM` icon (Figure A1.1). Note that just below the toolbar there is a status display which is initially green and displaying 'Idle'. Later it will change colour and indicate the current operation. Below the status display are four buttons for use when modelling, especially when fitting the ethnicity model to three-dimensional photogrammetric data (Section A1.2.1).

The `File` menu has the following menu items:

- `Edit Attributes`. This is for development work, and not users. It is used during editing of the modeller's target shapes.
- `New Model`. A new model is named and its mesh is displayed, ready to be modified parametrically (see below). The initial shape of the mesh is generic and only a starting point for creating a face model of specific shape and proportions.
- `Edit Model`. Selects a specific model from the listing of all existing models. The current model name is displayed instead of 'Idle' in the status line.
- `Save Model`. This saves the set of attribute-value pairs for the model currently being edited. They are saved to the associated `.json` file. That file has the same name as the model, with suffix `.json`. This option is greyed out unless a model is open for editing.

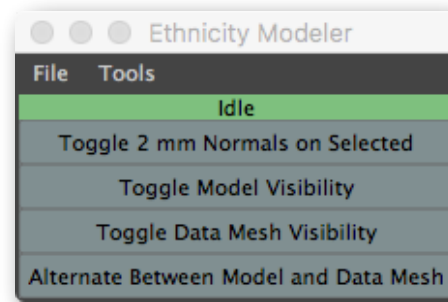


Figure A1.1. The initial state of the Ethnicity Modeller user interface.

- `Save Model As`. This allows the currently-open model to be saved with a new name, allowing one model to be the basis for another. This option is greyed out unless a model is open for editing. The name of that model changes in the status display to this new name.
- `Save Smooth Mesh`. This exports an OBJ format mesh of the current model. (It is greyed out unless a model is open for editing.)
- `Close Model`. This will prompt to save first, close without saving, or cancel. This menu option is greyed out unless a model is open for editing.)

The `Tools` menu is intended for visualisation once models have been created. The first is 'Interpolate A-B' which allows the user to select two models whereupon a set of window with sliders to control the interpolation is created. The options are to interpolate all regions simultaneously, or to blend various regions individually. For instance, if one model were `EAS_M20`, the average of 20 East Asian males, and the other model were `EUR_M20`, the European counterpart, then all attributes could be interpolated in parallel, or perhaps only those of the eyes or nose interpolated.

## A1.2 Model Editing

With a model selected and open for editing, a multi-tab window is launched, with tabs for adjusting the Cranium, Eye, Face, Jaw, Mouth, and Nose (see Figure A1.2). Each tab organises the attributes and current values for that region. See Section 5.3 regarding the terminology used for naming the attributes and their description. Each facial attribute is adjusted by a corresponding slider that varies over a normalised range from -1.0 to 1.0, or

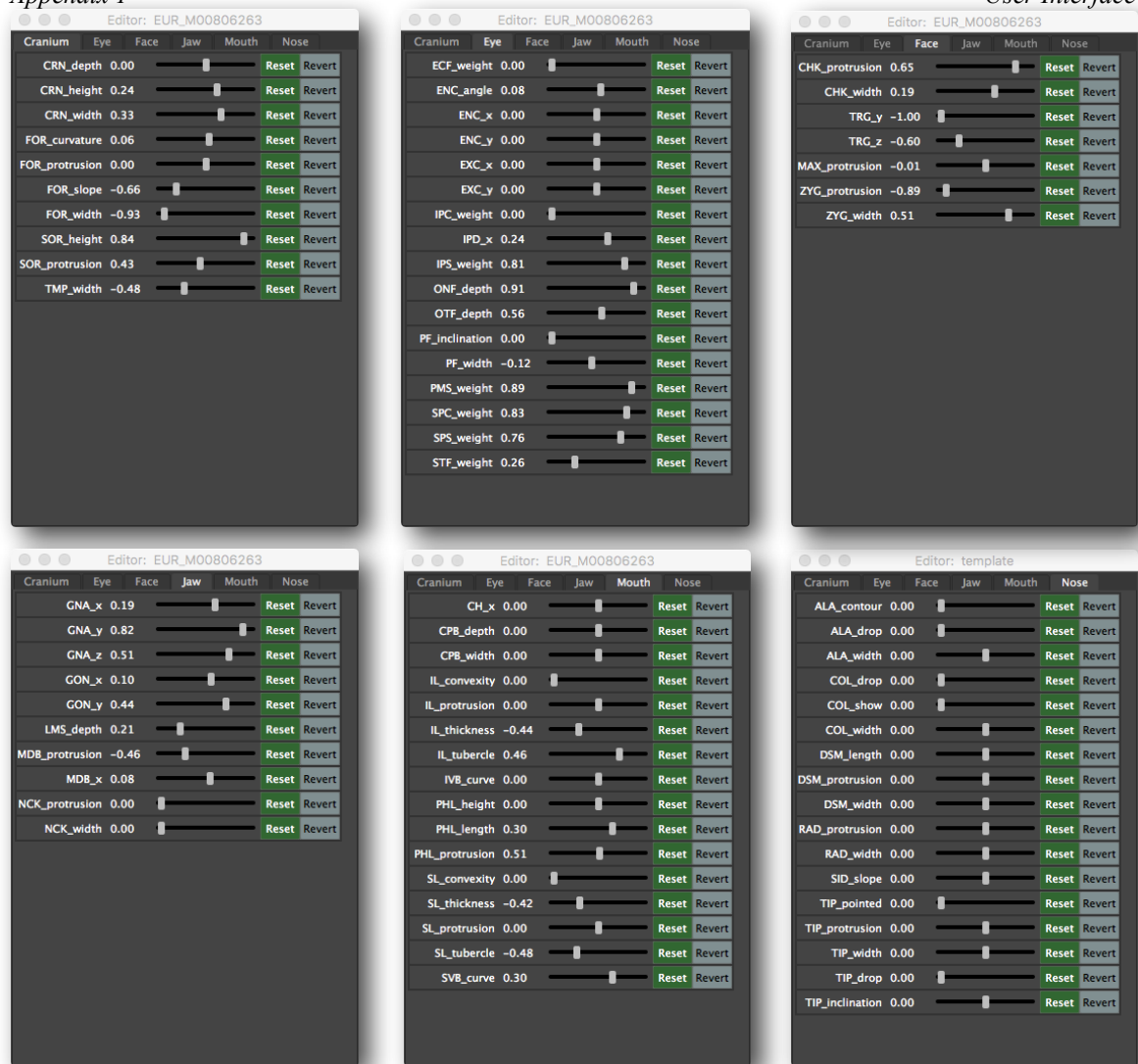


Figure A1.2. The attribute editor consists of six tab panes. This shows the attributes for the eye, and their current values, with buttons to reset to zero and revert to last saved.

from 0.0 to 1.0, depending on the attribute. The meaning of a specific attribute is often best appreciated by observing its effect on the face model as the corresponding parameter is adjusted in the user interface.

Each facial region has multiple features, each identified by a capitalised abbreviation such as COL for the columella of the nose region, or GNA for the gnathion of the jaw region. Those attributes that directly relate to the position or placement of a landmark operate in only one dimension within a Cartesian coordinate system with X positive to the model's left, Y positive upward, and Z positive ahead of the model. For example, GNA\_y corresponds to the vertical position of the chin and CH\_x corresponds to the hor-

horizontal position of the cheilion (the corner of the mouth). Attributes associated with dimensions rather than positions are identified with either ‘width’ for lateral (X) dimensions, ‘length’ or ‘height’ with vertical dimensions (Y), and ‘protrusion’ and ‘depth’ with Z. Sulci, folds, contours, curves and convexities are defined using ‘weight’ to indicate the degree of definition or magnitude of the feature.

The additivity associated with combinations of attributes can be explored interactively. Unintended distortion have been minimised but not entirely avoided (Section 5.4). Also, some combinations of attributes would not co-occur in extreme combinations, but individually would occur in one or another extreme.

#### *A1.2.1 Attribute Dependency*

The six facial regions (cranium, eyes, face, jaw, mouth, and nose) are thus arranged about the model origin, and will be deformed to ‘match’ (see Section 5.6 regarding criteria) the geometry of the model mesh to the surface geometry of a given face that is the subject of modelling. Fortunately, the features of the cranium are above and posterior to the model origin, and vary independently of those of the other five regions. The other five regions, however, all have significant dependencies on one specific attribute: the degree of protrusion of the maxilla. This attribute establishes the overall protrusion (in Z) of the cheeks and the base of the nose and the mouth, since (as expected anatomically) those other facial regions ‘ride’ upon the maxilla. It is therefore recommended that, as the first step in modelling, to adjust `MAX_protrusion` (in the Face tab). Specifically, the attribute `MAX_protrusion` is adjusted to set the alar base (the posterior of the alar wings where the nose blends with the adjacent cheeks) to the right degree of protrusion in Z relative to the eyes for the given ethnicity being modelled.

The attribute `MAX_protrusion` is an important facial attribute for capturing ethnic variation, of course, as it characterises prognathism (in addition to independent protrusion of the mandible and gnathion, i.e., the attributes `MDB_protrusion` and `GNA_z`).

With `MAX_protrusion` set, the nose, mouth, cheeks, and jaw attributes can then be adjusted relative to the base of the nose, and likewise those of the mouth, cheeks, and jaw. If `MAX_protrusion` were subsequently changed, however, the attributes of those other regions would have to be re-adjusted, hence it should be set first. With `MAX_protrusion` determined, the attributes of other regions can then be adjusted with minimal order dependency.

To create a new model, select from the user interface `File>New Model`, then provide a unique model name. A new file will be initialised with default values for all attribute-value pairs, and written as a JSON file. The attributes are then adjusted to fit the model geometry to the given subject face, having first established `MAX_protrusion`, as discussed. Finally, the model is written to its JSON file prior to closing the user interface.

#### *A1.2.2 Matching to Photogrammetric Data*

If a model exists for which a three-dimensional photogrammetric data mesh is also available, it will be loaded automatically and placed in the layer `Data_Mesh_Layer`.

Preparation of the photogrammetric data mesh (e.g., in the form of a textured obj file) requires placing that data mesh be scaled, rotated, and translated in order to align it with the model mesh. The corneas in the data mesh are aligned to closely superimpose upon those of the model mesh (to achieve this, model's `IPD_x` attribute is adjusted to match the X spacing between the eyes of the subject data). With the data mesh placed so that (like those of the model), the two corneas just touch the X axis and straddle the origin ( $Y = Z = 0.0$ ), the pivot for rotation of the data mesh is then set to the origin (so that the data mesh would rotate about the X axis, effectively pivoting the entire data mesh about the corneas, keeping the eyes in registration with the model, but adjusting for orientation in the ZY plane). There is, therefore, one additional degree of freedom to adjust: rotation of that mesh about the X axis in order to align the Frankfort plane in the data mesh with the XZ plane of the model.

With the photogrammetric data pre-aligned with the model, that data is imported automatically when the model is loaded. Provided hardware texturing is enabled, the data mesh, placed in the `Data_Mesh_Layer`, will be able to be alternately superimposed or hidden. The visibility of a data mesh, and that of the model, can be controlled by the three buttons on the main UI:

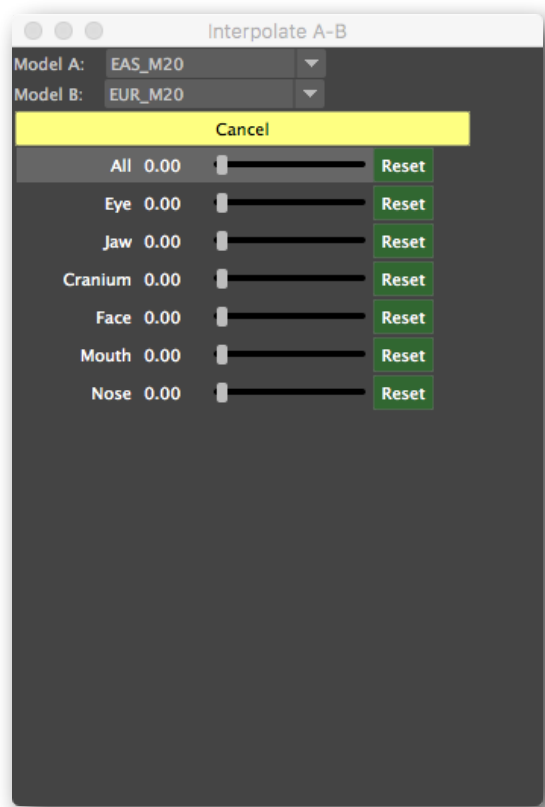
- `Toggle Model Visibility`. This toggles the visibility of the model, which is placed in the `Render_Layer`.
- `Toggle Data Mesh Visibility`. This toggles the visibility of the photogrammetric mesh, which is placed in the `Data_Mesh_Layer`.
- `Alternate Between Model and Data Mesh`. The two meshes alternate visibility in order to compare the model to the subject data.

First, the base of the nose in the model mesh is translated in Z by `MAX_protrusion` until in rough alignment with the corresponding area of the data mesh, alternating visibility of the data mesh as necessary. Final adjustment is assisted by selecting either the model mesh or the data mesh and clicking `Toggle 2mm Normals on Selected` (Figure A1.1). This permits visual inspection of the local displacement between the two meshes. Two millimetres is chosen as that distance is less than one standard deviation in the distribution of facial dimensions. Some features such as the shape of the lips and eyelids are more subtle, of course.

### A1.3 Ethnic Difference Visualisation

After having created a set of models corresponding to individuals or averages of various ethnicities, one of the immediate benefits of a parametric modelling system is the ability to continuously vary the values of its attributes in order to produce animated interpolation (see Section 2.3.3).

In the user interface, the menu item `tools>Interpolate A-B` launches an interface in which two models can be specified, then the model mesh can display any linear com-



*Figure A1.3. A tool for interpolating between two models, in this case EAS\_M20 and EUR\_M20, the averages of 20 male individuals each of EAS and EUR ethnicity. The model can either be interpolated with all attributes blending linearly, or on a per region basis independently.*

combination of the two, either by interpolating all attributes simultaneously across all regions, or on a per-region basis.



## Appendix 2. Ethnicity Modeller Code

The Ethnicity Modeller, or *EM*, is written in Python 3.5 on Maya 2017. It consists of the following script files:

<code>config.py</code>	definition of system constants and paths
<code>launchUI.py</code>	start the user interface
<code>Model.py</code>	instantiates and edits model, modify attributes
<code>ModellerUI.py</code>	user interface (widgets and callback functions)
<code>ModelExporter.py</code>	exports .json files to a spreadsheet in CSV format

In terms of operation, a tool is added to a shelf to launch the user interface. When launched, a window is created with the controls for the Ethnicity Modeller (see Appendix 1). The layout of the user interface is specified in `ModellerUI.py`, along with the callback functions which perform the operations associated with the various pulldown menu items and sliders. These callback functions in turn call functions in `Model.py`, which result in operations in Maya, including the setting of blendshape target weights. A stand-alone file `ModelExporter.py` is provided to create a spreadsheet in CSV format that assembles all the model files (in JSON format) that have been created and stored. The code is presented in the following sections.

### Appendix 2.1 The Configuration File

```
# config.py
import os
import userConfig

CSV_PATH      = os.path.join(userConfig.ROOT_PATH, 'data', 'csv')
MESH_PATH     = os.path.join(userConfig.ROOT_PATH, 'data', 'obj')
MODEL_PATH    = os.path.join(userConfig.ROOT_PATH, 'data', 'json')
DATA_MESH_PATH = os.path.join(userConfig.ROOT_PATH, 'data', 'photogrammetry')

ALL = 'All'
IDLE_STATE      = 'idle'
VIEWING_STATE   = 'viewing'
MODIFIED_STATE  = 'modified'
INTERPOLATING_STATE = 'interpolate'

# this base mesh has a pre-made blend shape (base_blendShape) all ready
BASE_FOR_EDITING = 'base'
MODEL_LAYER      = 'Render_Layer'
```

```

# this layer is used for displaying photogrammetric examples
DATA_MESH_LAYER = 'DataMesh_Layer'

# thus far no reason not to re-use base for interpolation as well
BASE_FOR_INTERPOLATION = 'base'
INTERPOLATION_LAYER = 'Render_Layer'

# the eyes are translated in x according to attribute IPD_x, within the
# specified range
IPD_X_MIN = 25.5
IPD_X_MED = 30
IPD_X_MAX = 35

# some standardized colors
BEIGE = [0.5, 0.75, 0.5]
DARK_GRAY = [0.2, 0.2, 0.2]
DARK_GREEN = [0.2, 0.4, 0.2]
DARK_MOAVE = [0.22, 0.39, 0.39]
DEFAULT_GRAY = [0.267, 0.267, 0.267]
GRAY_GREEN = [0.1, 0.3, 0.1]
LIGHT_GRAY = [0.8, 0.8, 0.8]
LIGHT_GREEN = [0.5, 0.85, 0.57]
MID_GRAY = [0.4, 0.4, 0.4]
MOAVE = [0.5, 0.56, 0.57]
RED = [1.0, 0.0, 0.0]
TAN = [0.9, 0.7, 0.56]
YELLOW = [1.0, 1.0, 0.53]

```

## Appendix 2.2 The User Interface Launcher

```

#launchUI.py

import ModellerUI as ui
import userConfig as uc
import config as c
import Model as m

reload(ui)
reload(uc)
reload(c)
reload(m)

interface = ui.ModellerUI()
interface.startUI()

```

## Appendix 2.3 The Three-Dimensional Model

```

## Model.py

import maya.cmds as cmds
import maya.mel as mel

import json
import glob

```

```

from config import *

#
class Model():
    """ A Model consists of a base (i.e., a deformable polygonal mesh), and an
    associated blend shape deformer. All models involve the same mesh
    topology so that they are morphable. A model instance in a Maya scene
    is a temporary binding between the base, its associated blend shape
    node, and some instance variables (name, dir, etc.). The initializer
    can reference any base. If no blendShape is specified, it is assumed
    already in the history of that base. Note that the base, being a
    polygonal mesh, can be loaded elsewhere from an OBJ file, and treated
    as a target for a blendShape."""

#
    def __init__(self, name, base, blendShape=None):
        """ This constructs a model instance which references a polygonal mesh,
        (base) with its associated blendShape. The Model instance is a
        binding between this base, blend shape, and manages the loading and
        saving of its target weights. """

        # the name of the model can be changed by saveModelAs
        self._name = name

        # base is the specified polygonal mesh to be deformed by the blendshape
        self._base = base

        # the base mesh is presumed to already have an associated blend shape
        # so either use the blendshape specified by the kwarg or find it.
        if blendShape:
            self._blendshape = blendShape
        else:
            # find the blendshape node for this base from its history
            history = cmds.listHistory(self._base)
            blendShapes = []
            for obj in history:
                if cmds.objectType(obj) == 'blendShape':
                    blendShapes.append(obj)
            if len(blendShapes) == 0:
                print('no blendShape found for base = %s' % base)
            if len(blendShapes) > 1:
                print('multiple blendShapes found for base = %s: %s'%
                    (base, blendShapes))
            self._blendShape = blendShapes[0]

        # Model data is read, modified, and saved based on the list self._data.
        # Since all models (each a JSON file) should be variations on the same
        # template (and differing only in the attribute values) the template is
        # first read in, then each attribute in the template is given the value
        # from the model if available. A new attribute can also be introduced
        # for which the model had no prior value, and likewise, an attribute in
        # the model that is no longer in the template can be deprecated by
        # dropping it from the template. The template data now is written out
        # to become the new version of the model. The variable self._data is
        # written when the model is saved. Quick access to it is provided by
        # a dictionary self._dict, with the attribute name as key and the
        # corresponding dict for that attribute as the associated value.
        self._dict = {}
        self._modified = {}
        self._numberModified = 0

```

```

# we start with loading up self._data from the template, and populate
# a dictionary self._dict to map between the name of an attribute and
# its associated dictionary. At the same time, another dictionary
# self._modified will indicate if a given attribute has been modified

self._data = self.readModel(name='template')
for dict in self._data:
    attribute = dict['attribute']
    self._dict[attribute] = dict
    self._modified[attribute] = False

# next, import the current model data (if the model is new, the default
# values in the template are used). Note that only those attributes
# that are found in the current template are assigned values from the
# current model.
if self._modelExists(name):
    currentData = self.readModel(name)
    for currentDict in currentData:
        attribute = currentDict['attribute']
        if attribute in self._dict:
            # then copy its value over from the template version
            self._dict[attribute]['value'] = currentDict['value']

# now save the refreshed data back to JSON
self.saveModel()

# update the targets according to the attribute values
# and save a copy of the initial value of each attribute to allow
# reverting to saved.
self._initialValue = {}
# set up a dictionary to quickly give the min value for each attribute,
self._minValue = {}
for dict in self._data:
    attribute = dict['attribute']
    value = dict['value']
    self._initialValue[attribute] = value
    self._minValue[attribute] = dict['minValue']
    self.setAttributeValue(attribute, value)

#=====
#                                     S E T T E R S   /   G E T T E R
#
#
#-----
def getModelPath(self, name=None):
    """ returns a path consisting of the default model directory path with
        the name (or default self._name) and adds the suffix '.json'."""

    name = name if name else self._name
    return MODEL_PATH + '/' + name + '.json'

#=====
#
#                                     P U B L I C   F U N C T I O N
#
#-----
def modelModified(self):
    """ Returns true if the model's current values are not the same as when
        initially read from the JSON file. """

    return self._numberModified > 0

```

```

#
def readModel(self, name):
    """ Returns the specified JSON file. """

    with open(self.getModelPath(name=name)) as f:
        data = json.load(f)
    return data

#
def saveModel(self):
    """ Writes all blendShape target weights (model attributes) into a JSON
    file. The JSON file has supplemental information for each dict,
    such as the region and a string for annotating the description.
    Hence saving involves first reading in a template for
    that fixed information, then overwriting the default (0.0) weight
    values for only those attribute values represented by blend shape
    target weights. """

    with open(self.getModelPath(), 'w') as f:
        json.dump(self._data, f, indent=4)

#
def saveModelAs(self, name):
    """ Changes the model's name and saves data with that new name. """

    self._name = name
    self.saveModel()

#
def attributeModified(self, attribute):
    """ returns whether the value of this attribute has been modified. """

    return self._modified[attribute]

#
def setLayer(self, layer):
    cmds.editDisplayLayerMembers(layer, self._base)

#
def setLayerVisibility(self, layer, on=True):
    cmds.setAttr('%s.visibility' % layer, on)

#=====
#
#
#
#
def getMeshPath(self, name=None):
    """ returns a path consisting of the default mesh directory path with
    the name (or default self._name) and adds the suffix '.obj'. """

    name = name if name else self._name
    return MESH_PATH + '/' + name + '.obj'

#
def saveMesh(self, fileName=None, divisions=0):
    """ Exports as an OBJ file a duplicate of the base for this model. The
    file name defaults to the model name. If divisions is 0 only the
    the control vertex of the base mesh are written, while divisions of
    2 (or more) results in creation of an OBJ file of a smooth,
    higher-polygon count mesh. """

```





```

        self.setAttributeValue(attribute, value)
        return value

#
def getAttributeMinValue(self, attribute):
    """ returns the minimum value of the attribute. """

    return self._dict[attribute]['minValue']

#
def getAttributeMaxValue(self, attribute):
    """ returns the minimum value of the attribute. """

    return self._dict[attribute]['maxValue']

```

## Appendix 2.4 The User Interface

```

# ModellerUI.py

import maya.cmds as cmds
import maya.mel as mel
import glob
from functools import partial

import Model as mm
from config import *

#
class ModellerUI(object):
    """ The ModellerUI is responsible for managing and displaying a model.
        A Model instance contains a reference to a base mesh and a blendShape,
        both in the Maya scene. All meshes have the same topology, so that any
        mesh can serve as either a base or a target (or both).

        To control this generality, a few basic tools are provided. The
        first is the model editor, which uses the full set of targets to morph
        the base mesh. The weights can then be saved as in a JSON format, and
        the mesh itself can be saved as an OBJ file (so that it has 'baked in'
        the deformations). OBJ files are quick loading, as are JSON files for
        a blend shape that is already set up with targets. Creating a blend
        shape with targets is very slow, however, so the Modeller scene has one
        prepared ahead of time (base_blendShape).

        Loading operations:
        1) assigning a set of target weights to a pre-existing blend shape for
        a given mesh, loaded in from a JSON file.
        2) loading a mesh as an OBJ file and naming it in the scene.

        Saving operations:
        1) saving the complete set of target weights to a JSON file. It is
        assumed there is a 1:1 correspondence between attribute names and
        blend target.
        2) saving a mesh as an OBJ file with new name. """

#
def __init__(self):
    """ A user interface is constructed from a window with a pulldown menu
        and a layout which is modified according to which tool is in
        operation. The File menu has options that are grayed out when not

```



```

        applicable, so references are saved for these menu items for later
        editing. The interface communicates with the Maya model via the
        creation of Model instances. """
self._state = None
self._mainWindow = None
self._modelText = None
self._modelEditor = None
self._newModelMI = None
self._editModelMI = None
self._saveModelMI = None
self._saveModelAsMI = None
self._saveMeshMI = None
self._saveSmoothMeshMI = None
self._closeModelMI = None

# required to associated a slider with a given attribute
self._attributeSliderDict = None

# a model instance and the
self._model = None
self._dataMesh = None

# the name of the model (which is the name of a JSON file)
self._modelName = None

#
def startUI(self):
    """ Launch the UI. """

    # remove any old Modeller UIs
    currentWindows = cmds.lsUI(windows=True)
    for window in currentWindows:
        title = cmds.window(window, query=True, title=True)
        if title == 'Ethnicity Modeller':
            cmds.deleteUI(window)
        elif title.startswith('Editor'):
            cmds.deleteUI(window)
        elif title.startswith('Interpolator'):
            cmds.deleteUI(window)

    self._mainWindow = cmds.window(
        title='Ethnicity Modeller',
        width=200,
        menuBar=True,
        maximizeButton=False)

    self._fileMenu = cmds.menu(label='File')

    # this edit the targets
    self._editAttributesMI = cmds.menuItem(
        label='Edit Attributes',
        command=lambda *args: self._editAttributesCB())
    # this will create a new model from the generic (full attributes) model
    self._newModelMI = cmds.menuItem(
        label='New Model',
        command=lambda *args: self._newModelCB())

    # edit an existing model, based on the name selected from a submenu
    self._editModelMI = cmds.menuItem(
        label='Edit Model',
        subMenu=True,

```

```

        postMenuCommand=lambda *args: self._editModelMenuCB())
cmds.setParent(self._fileMenu, menu=True)

# save the currently open Model
self._saveModelMI = cmds.menuItem(
    label='Save Model', command=lambda *args: self._saveModelCB())

# save the currently open model to a new model name
self._saveModelAsMI = cmds.menuItem(
    label='Save Model As...',
    command=lambda *args: self._saveModelAsCB())

# export the base mesh to obj
self._saveMeshMI = cmds.menuItem(
    label='Save Mesh', command=lambda *args: self._saveMeshCB())

# export a smooth (higher polygon count) mesh to obj
self._saveSmoothMeshMI = cmds.menuItem(
    label='Save Smooth Mesh',
    command=lambda *args: self._saveSmoothMeshCB())

cmds.menuItem(divider=True)

# close the current editing session
self._closeModelMI = cmds.menuItem(
    label='Close Model', command=lambda *args: self._closeEditorCB())

# these tools require the idle state
cmds.setParent('..')
cmds.menu(label='Tools', tearOff=True)

cmds.menuItem(
    label='Interpolate A-B',
    command=lambda *args: self._interpolateABToolCB())

cmds.menuItem(label='Model Comparator', enable=False)
cmds.menuItem(label='Individual Differences Visualizer', enable=False)

cmds.rowColumnLayout()
self._modelText = cmds.text(
    width=200,
    label='Idle',
    align='center',
    backgroundColor=BEIGE)

cmds.button(
    'Toggle 2 mm Normals on Selected',
    backgroundColor=MOAVE,
    command=lambda *args: self._toggleNormalsCB())

cmds.button(
    'Toggle Model Visibility',
    backgroundColor=MOAVE,
    command=lambda *args: self._toggleModelVisibilityCB())

cmds.button(
    'Toggle Data Mesh Visibility',
    backgroundColor=MOAVE,
    command=lambda *args: self._toggleExampleVisibilityCB())

cmds.button(

```

```

        'Alternate Between Model and Data Mesh',
        backgroundColor=MOAVE,
        command=lambda *args: self._alternateVisibilityCB())
cmds.showWindow()
self._setState(IDLE_STATE)

#
def _setState(self, state):
    """ Control of whether each menu is enabled or disabled. """

    self._state = state

    if self._state == IDLE_STATE:
        cmds.menuItem(self._editAttributesMI, edit=True, enable=True)
        cmds.menuItem(self._newModelMI, edit=True, enable=True)
        cmds.menuItem(self._editModelMI, edit=True, enable=True)
        cmds.menuItem(self._saveModelMI, edit=True, enable=False)
        cmds.menuItem(self._saveModelAsMI, edit=True, enable=False)
        cmds.menuItem(self._closeModelMI, edit=True, enable=False)
        cmds.menuItem(self._saveMeshMI, edit=True, enable=False)
        cmds.menuItem(self._saveSmoothMeshMI, edit=True, enable=False)
        cmds.text(
            self._modelText, edit=True, label='Idle',
            backgroundColor=BEIGE)

    elif self._state == VIEWING_STATE:
        cmds.menuItem(self._editAttributesMI, edit=True, enable=True)
        cmds.menuItem(self._newModelMI, edit=True, enable=True)
        cmds.menuItem(self._editModelMI, edit=True, enable=True)
        cmds.menuItem(self._saveModelMI, edit=True, enable=False)
        cmds.menuItem(self._saveModelAsMI, edit=True, enable=True)
        cmds.menuItem(self._closeModelMI, edit=True, enable=True)
        cmds.menuItem(self._saveMeshMI, edit=True, enable=True)
        cmds.menuItem(self._saveSmoothMeshMI, edit=True, enable=True)
        cmds.text(
            self._modelText,
            edit=True,
            label=format('Editing %s' % self._modelName),
            backgroundColor=TAN)

    elif self._state == MODIFIED_STATE:
        cmds.menuItem(self._editAttributesMI, edit=True, enable=False)
        cmds.menuItem(self._newModelMI, edit=True, enable=False)
        cmds.menuItem(self._editModelMI, edit=True, enable=False)
        cmds.menuItem(self._saveModelMI, edit=True, enable=True)
        cmds.menuItem(self._saveModelAsMI, edit=True, enable=True)
        cmds.menuItem(self._closeModelMI, edit=True, enable=True)
        cmds.menuItem(self._saveMeshMI, edit=True, enable=True)
        cmds.menuItem(self._saveSmoothMeshMI, edit=True, enable=True)
        cmds.text(
            self._modelText,
            edit=True,
            label=format('Editing %s' % self._modelName),
            backgroundColor=TAN)

    elif self._state == INTERPOLATING_STATE:
        cmds.menuItem(self._editAttributesMI, edit=True, enable=False)
        cmds.menuItem(self._newModelMI, edit=True, enable=False)
        cmds.menuItem(self._editModelMI, edit=True, enable=False)
        cmds.menuItem(self._saveModelMI, edit=True, enable=False)
        cmds.menuItem(self._saveModelAsMI, edit=True, enable=True)

```

```

        cmds.menuItem(self._closeModelMI,      edit=True, enable=False)
        cmds.menuItem(self._saveMeshMI,        edit=True, enable=False)
        cmds.menuItem(self._saveSmoothMeshMI,  edit=True, enable=False)
        cmds.text(
            self._modelText,
            edit=True,
            label=format('Interpolating'),
            backgroundColor=TAN)

#
def _editAttributesCB(self):
    """ Used to edit targets. """

    self._createModelEditorCB('template', False)

#
def _newModelCB(self):
    """ Asks for a new model name, then off to _createModelEditor. """

    result = cmds.promptDialog(
        title='Create New Model',
        message='Enter Name:',
        button=['OK', 'Cancel'],
        defaultButton='OK',
        cancelButton='Cancel',
        dismissString='Cancel')
    if result == 'Cancel':
        return

    # fetch the name and make sure it is new
    name = cmds.promptDialog(query=True, text=True)
    if self._modelExists(name):
        # complain if the name is not new
        cmds.confirmDialog(
            backgroundColor=RED,
            messageAlign='center',
            title=format('A model named %s already exists.' % name),
            message='Please choose another name.',
            button=['OK'])
        return

    self._createModelEditorCB(name, True)

#
def _editModelMenuCB(self):
    """ Select an existing model name, then off to _createModelEditor. """

    # delete the previous model edit menu
    cmds.setParent(self._editModelMI, menu=True)
    cmds.menu(self._editModelMI, edit=True, deleteAllItems=True)

    # compile a list of the current models and create menu items for them
    models = self._getModelNames()
    self._modelsMenuItems = []
    for m in models:
        cmds.menuItem(
            label=m, command=partial(self._createModelEditorCB, m))

#
def _createModelEditorCB(self, name, x):
    """ Creates a model editor (a column of attribute sliders). """

```

```

self._modelName = name
self._model = mm.Model(name, BASE_FOR_EDITING)
modelData = self._model.readModel(name)

# delete the previous model editor, if any
currentWindows = cmds.lsUI(windows=True)
for window in currentWindows:
    title = cmds.window(window, query=True, title=True)
    if title.startswith('Editor'):
        cmds.deleteUI(window)

# just in case there is another model open (but not modified), close it
# to remove any associated dataMesh.
self._closeModel()

# place this editor window just beside the main window
tlc = cmds.window(self._mainWindow, query=True, topLeftCorner=True)
tlc[1] += 250

self._modelEditor = cmds.window(
    title=format('Editor: %s' % name),
    widthHeight=(350, 500),
    topLeftCorner=tlc,
    maximizeButton=False)

form = cmds.formLayout()

tabs = cmds.tabLayout(
    innerMarginWidth=4,
    innerMarginHeight=4,
    scrollable=True)
cmds.formLayout(
    form,
    edit=True,
    attachForm=((tabs, 'top', 0),
                (tabs, 'left', 0),
                (tabs, 'bottom', 0),
                (tabs, 'right', 0)))

# make a tab for each region
region = modelData[0]['region']

#tab = cmds.frameLayout(region, backgroundColor=DARK_GRAY)
tab = cmds.rowColumnLayout(backgroundColor=DARK_GRAY)
cmds.tabLayout( tabs, edit=True, tabLabel=(tab, region))

# initialize a dictionary for looking up slider based on attribute name
self._attributeSliderDict = {}

for attributeDict in modelData:
    if attributeDict['region'] != region:
        region = attributeDict['region']
        cmds.setParent('..')
        # tab = cmds.frameLayout(region, backgroundColor=DARK_GRAY)
        tab = cmds.rowColumnLayout(backgroundColor=DARK_GRAY)
        cmds.tabLayout(tabs, edit=True, tabLabel=(tab, region))
        self._createAttributeSlider(attributeDict)

cmds.showWindow(self._modelEditor)
self._setState(VIEWING_STATE)

```

```

        # if a corresponding photogrammetric example exists, load it
        self._loadDataMesh(self._modelName)

#
def _createAttributeSlider(self, attributeDict, step=0.05):
    """ Creates a slider for a given attribute (name), weight (value), and
        description (annotation), and min and max values
        for the slider. """

    attribute = attributeDict['attribute']
    value      = attributeDict['value']

    # put one slider group and reset button per row
    cmds.rowColumnLayout(numberOfRows=1)

    self._attributeSliderDict[attribute] = cmds.floatSliderGrp(
        label=attribute,
        columnWidth3=[85, 50, 100],
        columnAlign3=['right', 'center', 'center'],
        annotation= attributeDict['description'],
        field=True,
        min=attributeDict['minValue'],
        max=attributeDict['maxValue'],
        backgroundColor=DEFAULT_GRAY,
        step=step,
        sliderStep=step,
        value=value,
        changeCommand=partial(self._attributeSliderCB, attribute),
        dragCommand=partial(self._attributeSliderCB, attribute))

    cmds.button(
        label='Reset',
        width=40,
        backgroundColor=DARK_GREEN,
        command=partial(self._resetAttributeSliderCB, attribute))

    cmds.button(
        label='Revert',
        width=40,
        backgroundColor=MOAVE,
        command=partial(self._revertAttributeSliderCB, attribute))
    cmds.setParent('..')

#
def _attributeSliderCB(self, attribute, x):
    """ This callback takes the slider value for the given attribute and
        sets the weight of the corresponding target in the model. """

    slider = self._attributeSliderDict[attribute]
    value = cmds.floatSliderGrp(slider, query=True, value=True)
    self._setSlider(attribute, value)

#
# def _attributeSliderCB(self, attribute, x, intervals=7):
#     """ This callback takes the slider value for the given attribute and
#         sets the weight of the corresponding target in the model. """
#
#     slider = self._attributeSliderDict[attribute]
#     sliderValue = cmds.floatSliderGrp(slider, query=True, value=True)
#

```

```

#     nearestValue = round(sliderValue*intervals)/intervals
#     self._setSlider(attribute, nearestValue)

#
def _resetAttributeSliderCB(self, attribute, x):
    """ This callback resets to zero the given attribute. """

    self._setSlider(attribute, 0)

#
def _revertAttributeSliderCB(self, attribute, x):
    """ This resets to attribute's slider to its initial value. """

    initialValue = self._model.getAttributeValue(
        attribute, initialValue=True)
    self._setSlider(attribute, initialValue)

#
def _setSlider(self, attribute, value):
    """ This sets a given slider to a given value, and sets the background
        color to either the default gray or a mid gray (mild highlight) if
        the slider is set to other than the attribute's initial value. """

    slider = self._attributeSliderDict[attribute]
    initialValue = self._model.getAttributeValue(
        attribute, initialValue=True)
    modified = value != initialValue

    cmds.floatSliderGrp(slider, edit=True, value=value)
    cmds.floatSliderGrp(
        slider,
        edit=True,
        backgroundColor=MID_GRAY if modified else DEFAULT_GRAY)

    # also actually set the model's value for that attribute!
    self._model.setAttributeValue(attribute, value)

    if self._model.modelModified() and self._modelName != 'template':
        self._setState(MODIFIED_STATE)
    else:
        self._setState(VIEWING_STATE)

#
def _getModelNames(self):
    """ Returns a list of all models (without the .json suffix). """

    paths = glob.glob(MODEL_PATH + "/*.json")
    models = []

    for p in paths:
        file = os.path.split(p)[1]
        model = file.split('.')[0]
        models.append(model)

    return sorted(models)

#
def _modelExists(self, name):
    """ Returns whether a model with that name already exists. """

    paths = glob.glob(MODEL_PATH + "/*.json")

```

```

        for path in paths:
            file = os.path.split(path)[1]
            modelName = file.split('.')[0]
            if modelName == name:
                return True
        return False

#
def _closeEditorCB(self):
    """ Prompts to ask to save if it has been modified. """

    if self._model.modelModified() and self._modelName != 'template':
        result = cmds.confirmDialog(
            backgroundColor=YELLOW,
            messageAlign='center',
            title=format('%s has been modified.' % self._modelName),
            message='Save Model Before Closing?',
            defaultButton='Cancel',
            dismissString='Cancel',
            button=['Save', 'Do Not Save', 'Cancel'])
        if result == 'Save':
            self._saveModelCB()
        elif result == 'Cancel':
            return
        self._closeModel()

#
def _toggleModelVisibilityCB(self):
    """ Toggles the state of visibility of the model (Base_Layer). """

    state = cmds.getAttr(MODEL_LAYER + '.visibility')
    cmds.setAttr(MODEL_LAYER + '.visibility', not state)

#
def _toggleExampleVisibilityCB(self):
    """ Toggles the state of visibility of the example. """

    state = cmds.getAttr(DATA_MESH_LAYER + '.visibility')
    cmds.setAttr(DATA_MESH_LAYER + '.visibility', not state)

#
def _alternateVisibilityCB(self):
    """ Alternates between showing model only or data mesh only. """

    state = cmds.getAttr(MODEL_LAYER + '.visibility')

    cmds.setAttr(MODEL_LAYER + '.visibility', not state)
    cmds.setAttr(DATA_MESH_LAYER + '.visibility', state)

#
def _toggleNormalsCB(self):
    """ Toggles the state of displaying normals for whichever mesh is
        selected. """

    selection = cmds.ls(selection=True)
    if not selection:
        return

    for object in selection:
        #shape = cmds.listRelatives(selection, shapes=True)[0]

```



```

        shape = cmds.listRelatives(object, shapes=True)[0]

        # toggles the current state of display
        state = cmds.getAttr(shape + '.displayNormal')
        cmds.setAttr(shape + '.displayNormal', not state)

        # set normals to be face normals
        cmds.setAttr(shape + '.normalType', 1)

        # for 2 mm normals, set normalSize to 0.025 (or 0.035 for 3 mm)
        cmds.setAttr(shape + '.normalSize', 0.2)

#
def _closeModel(self):
    """ Cleanup to close a model. """

    # delete the previous attribute editor, if any
    if self._modelEditor:
        if cmds.window(self._modelEditor, query=True, exists=True):
            cmds.deleteUI(self._modelEditor)

    # remove any dataMesh and its material
    if self._dataMesh:
        mesh = cmds.ls(self._dataMesh)
        if mesh:
            cmds.delete(mesh[0])
            mel.eval('MLdeleteUnused;')
            self._dataMesh = None

    # update which menu items in the File menu are enabled versus disabled
    self._setState(IDLE_STATE)

#
def _saveModelCB(self):
    """ Saves the model with the current name, but NOT the template. """

    if self._modelName == 'template':
        cmds.confirmDialog(
            backgroundColor=[1.0, 0.0, 0.0],
            messageAlign='center',
            title='Sorry, the template cannot be modified.',
            button=['OK'])
        return
    self._model.saveModel()

    # refresh the model so it no longer is modified
    self._model = mm.Model(self._modelName, BASE_FOR_EDITING)
    self._setState(VIEWING_STATE)

#
def _saveModelAsCB(self):
    """ Prompts for a new name for the model data, then writes the model
        as a new JSON file. This allows saving the template as a new
        model. """

    # ask for a name for the new model
    result = cmds.promptDialog(
        title='Save Model As...',
        message='Enter Name:',
        button=['OK', 'Cancel'],
        defaultButton='OK',

```

```

        cancelButton='Cancel',
        dismissString='Cancel')

    if result == 'Cancel':
        return

    name = cmds.promptDialog(query=True, text=True)
    self._name = name
    cmds.text(self._modelText, edit=True, label=name)

    if self._modelEditor:
        title = format('Editor: %s' % name)
        cmds.window(self._modelEditor, edit=True, title=title)

    # have the model do the actual saving of the json file
    self._model.saveModelAs(name)

#=====
#                                     I N T E R P O L A T I O N   F U N C T I O N S
#
#
def _interpolateABToolCB(self):
    """ Prompts for two models A and B, then interpolates between them. """

    # delete the previous interpolation window, if any
    currentWindows = cmds.lsUI(windows=True)
    for window in currentWindows:
        title = cmds.window(window, query=True, title=True)
        if title.startswith('Interpolate A-B'):
            cmds.deleteUI(window)

    self._interpolatorWindow = cmds.window(
        width=300, maximizeButton=False, title='Interpolate A-B')

    cmds.columnLayout()

    # prompt to select model A
    self._modelA_optionMenu = cmds.optionMenu(label='Model A: ')

    cmds.menuItem('select')
    models = self._getModelNames()
    self._modelsMenuItems = []
    for m in models:
        cmds.menuItem(label=m)

    # prompt to select model B
    self._modelB_optionMenu = cmds.optionMenu(label='Model B: ')

    cmds.menuItem('select')
    models = self._getModelNames()
    self._modelsMenuItems = []
    for m in models:
        cmds.menuItem(label=m)

    # add a green button to start interpolation; it's action is to add the
    # interpolation sliders just below it in same window.
    self._startInterpolationButton = cmds.button(
        label='Interpolate',
        width=300,
        backgroundColor=[0.0, 1.0, 0.0],
        command=lambda *args: self._createInterpolationSlidersCB())

```

```

        cmds.showWindow()

#
def _cancelInterpolationCB(self):
    """ Kills any interpolator window. """

    currentWindows = cmds.lsUI(windows=True)
    for window in currentWindows:
        title = cmds.window(window, query=True, title=True)
        if title.startswith('Interpolate A-B'):
            cmds.deleteUI(window)
    self._setState(IDLE_STATE)

#
def _createInterpolationSlidersCB(self):
    """ Given two models, A and B, specified in _interpolateModels, this
        sets up the interpolation sliders. """

    # fetch the selection for model A
    modelA = cmds.optionMenu(
        self._modelA_optionMenu, query=True, value=True)

    # fetch the selection for model B
    modelB = cmds.optionMenu(
        self._modelB_optionMenu, query=True, value=True)

    # read the model data (JSON file) for model A
    self._modelName = modelA
    self._model = mm.Model(modelA, BASE_FOR_INTERPOLATION)
    modelAData = self._model.readModel(modelA)

    # read the model data (JSON file) for model B
    self._modelName = modelB
    self._model = mm.Model(modelB, BASE_FOR_INTERPOLATION)
    modelBData = self._model.readModel(modelB)

    # interpolation of per region is achieved by using a dictionary
    # of the attributes per region
    self._attributesPerRegion = {}

    # where associated with the key 'regionName' is a list of attrInterps,
    # each of which is a dictionary associating an attribute's name and the
    # values of that attribute for models A and B, plus the minValue which
    # is needed by the model's setAttributeValue method to determine which
    # target to set).

    # start with the first region in modelA (this assumes the models are in
    # sync)
    regionName = modelAData[0]['region']
    self._attrInterps = []

    for attrA, attrB in zip(modelAData, modelBData):
        # make a dictionary for this attribute
        attrInterp = {}
        attrInterp['attribute'] = attrA['attribute']
        attrInterp['valueA'] = attrA['value']
        attrInterp['valueB'] = attrB['value']
        attrInterp['minValue'] = attrA['minValue']
        # if region has changed, complete the entry for the previous region
        if attrA['region'] != regionName:

```

```

        self._attributesPerRegion[regionName] = self._attrInterps
        regionName = attrA['region']
        self._attrInterps = []
        self._attrInterps.append(attrInterp)

# and finish off the last region
self._attributesPerRegion[regionName] = self._attrInterps

# and add an entry for all regions
self._attributesPerRegion[ALL] = []

# make the slider for interpolating each region individually
cmds.setParent(self._interpolatorWindow)

# remove the button that started the interpolation
cmds.deleteUI(self._startInterpolationButton)

cmds.columnLayout()
# and replace it with a cancel button
cmds.button(
    'Cancel',
    width=300,
    backgroundColor=YELLOW,
    command=lambda *args: self._cancelInterpolationCB())

# initialize a dictionary for looking up slider based on attribute name
self._interpolationSliderDict = {}

# then make sliders for each region
for regionName in self._attributesPerRegion:
    self._createInterpolationSlider(regionName)

cmds.showWindow(self._modelEditor)

self._setState(INTERPOLATING_STATE)

#
def _createInterpolationSlider(self, regionName, step=0.05):
    """ Creates a slider for a given region, for alpha from 0 to 1. """

    # put one slider group and its associated reset button per row
    cmds.rowColumnLayout(numberOfRows=1)

    color = [0.4, 0.4, 0.4] if regionName == ALL else [0.267, 0.267, 0.267]

    self._interpolationSliderDict[regionName] = cmds.floatSliderGrp(
        label=regionName,
        columnWidth3=[75, 50, 100],
        columnAlign3=['right', 'center', 'center'],
        field=True,
        min=0,
        max=1,
        step=step,
        value=0,
        backgroundColor=color,
        changeCommand=partial(self._interpolationSliderCB, regionName),
        dragCommand=partial(self._interpolationSliderCB, regionName))

    # add a reset button for each region's slider
    cmds.button(
        label='Reset',

```

```

        width=40,
        backgroundColor=DARK_GREEN,
        command=partial(self._resetInterpolationSliderCB, regionName))
cmds.setParent('..')

# start off with alpha = 0
self._interpolateRegion(regionName, 0)

#
def _resetInterpolationSliderCB(self, regionName, x):
    """ This callback resets to zero the alpha value for the given
        interpolation slider for this region, which then sets all the
        target values to those of model A. """

    if regionName == ALL:
        # enable the 'All' slider if it is reset
        slider = self._interpolationSliderDict[regionName]
        cmds.floatSliderGrp(slider, edit=True, enable=True)

        for regionName in self._attributesPerRegion:
            slider = self._interpolationSliderDict[regionName]
            cmds.floatSliderGrp(slider, edit=True, value=0)
            self._interpolateRegion(regionName, 0)
        return

    # disable the 'All' slider if any other slider is reset
    slider = self._interpolationSliderDict[ALL]
    cmds.floatSliderGrp(slider, edit=True, enable=False)

    slider = self._interpolationSliderDict[regionName]
    cmds.floatSliderGrp(slider, edit=True, value=0)
    self._interpolateRegion(regionName, 0)

#
def _interpolationSliderCB(self, regionName, x):
    """ This callback takes the slider value for the given region and sets
        the weight of the corresponding targets in the model in that region
        to an interpolation according to the alpha value of the slider. """

    slider = self._interpolationSliderDict[regionName]
    alpha = cmds.floatSliderGrp(slider, query=True, value=True)

    if regionName == ALL:
        for regionName in self._attributesPerRegion:
            slider = self._interpolationSliderDict[regionName]
            cmds.floatSliderGrp(slider, edit=True, value=alpha)
            self._interpolateRegion(regionName, alpha)
        return

    # disable the 'All' slider if any other slider is moved
    slider = self._interpolationSliderDict[ALL]
    cmds.floatSliderGrp(slider, edit=True, enable=False)

    slider = self._interpolationSliderDict[regionName]
    cmds.floatSliderGrp(slider, edit=True, enable=True)
    self._interpolateRegion(regionName, alpha)

#
def _interpolateRegion(self, regionName, alpha):
    """ Interpolate all the attributes for this region. """

```

```

attrInterps = self._attributesPerRegion[regionName]
# iterate through all of the attributes for this region, determining
# the value as a LERP between A and B by alpha, and set the attribute
# accordingly
for attrInterp in attrInterps:
    attribute = attrInterp['attribute']
    valueA    = attrInterp['valueA']
    valueB    = attrInterp['valueB']
    value     = valueA + alpha*(valueB - valueA)
    self._model.setAttributeValue(attribute, value)

#=====
#                                     P H O T O G R A M M E T R I C   D A T A   F U N C T I O N S
#
#
#
def _loadDataMesh(self, name):
    """ This brings a 3D photogrammetric example into the Example_Layer,
        ready to use as reference while modeling. """

    # remove any serialization suffix, such as EAS_M20#2
    coreName = name.split('#')[0]
    # first see if this object is already loaded
    obj = cmds.ls(coreName)
    # delete the old object
    if obj:
        cmds.delete(obj)

    # and it's shader nodes
    mel.eval('MLdeleteUnused;')

    # now load the data mesh, if one exists
    if not self.dataMeshExists(coreName):
        return

    path = self._getDataMeshPath(coreName)
    cmds.file(path, i=True, type="mayaAscii", rpr=coreName)

    # place the object that was just imported into the data mesh layer
    transforms = cmds.ls(coreName + '*', type='transform')
    if transforms:
        self._dataMesh = transforms[0]
        cmds.editDisplayLayerMembers(DATA_MESH_LAYER, self._dataMesh)

#
def _getDataMeshPath(self, name):
    """ returns a path consisting of the default mesh directory path with
        the name (or default self._name) and adds the suffix '.obj'. """

    return DATA_MESH_PATH + '/' + name + '.ma'

#
def dataMeshExists(self, name):
    """ Returns True if an example exists for the given name. """

    paths = glob.glob(DATA_MESH_PATH + "/*.ma")

    for path in paths:
        file = os.path.split(path)[1]
        n = file.split('.')[0]
        if n == name:
            return True

```

```

        return False

#=====
#                                     M E S H   F U N C T I O N S
#
#
def _saveMeshCB(self):
    self._model.saveMesh()

#
def _saveSmoothMeshCB(self):
    fileName = self._modelName + '_smooth'
    self._model.saveMesh(divisions=2, fileName=fileName)

#
def _getMeshNames(self):
    """ Returns a list of all meshes (without the .obj suffix). """

    paths = glob.glob(MESH_PATH + "/*.obj")
    meshes = []

    for p in paths:
        file = os.path.split(p)[1]
        model = file.split('.')[0]
        meshes.append(model)
    return sorted(meshes)

```

## Appendix 2.5 The CSV Exporter

```

# ModelExporter.py

import os
import glob
import pandas as pd

class ModelExporter(object):
    DROPBOX_CSV_PATH = '/Dropbox/Neng/Maya/EthnicityModeller/data/csv'
    DROPBOX_MODEL_PATH = '/Dropbox/Neng/Maya/EthnicityModeller/data/json'
    user = 'sawitree'

    CSV_PATH = '/Users/' + user + DROPBOX_CSV_PATH
    MODEL_PATH = '/Users/' + user + DROPBOX_MODEL_PATH

    def getModelNames(self):
        """ Returns a list of all models (without the .json suffix). """

        paths = glob.glob(self.MODEL_PATH + "/*.json")
        models = []
        for p in paths:
            file = os.path.split(p)[1]
            model = file.split('.')[0]
            models.append(model)
        return sorted(models)

    def exportModelsToCSV(self):
        """ Creates a data frame from all models and exports it as CSV. """

        # compile a list of the current models
        models = self.getModelNames()

```

```
# base the data frame on the template
df = pd.read_json(self.MODEL_PATH + '/template.json')
for model in models:
    if model == 'template':
        continue
    d = pd.read_json(self.MODEL_PATH + '/' + model + '.json')
    df[model] = d['value']
df.to_csv(self.CSV_PATH + 'models.csv')

m = ModelExporter()
m.exportModelsToCSV()
```



## Appendix 3. Quantitative Analysis

### Appendix 3.1 Repeatability Study

In order to use the Ethnicity Modeller to create a specific model of a face, the adjustment of attribute values should be repeatable. That is, if the same face were modelled multiple times, the attribute values should have sufficiently small variation from trial to trial (where “sufficiently small” is discussed below).

An experiment was performed to study the repeatability of facial modelling, using two faces: the first is a three-dimensional data mesh derived from photogrammetry representing the average of 20 East Asian males (EAS\_M20 shown in Figure 6.24), and a second data mesh corresponding to the average mesh of 20 EUR males (EUR\_M20). Ten matching trials were performed by the author for each of the two meshes; in each trial the attributes of the *EM* were adjusted until the model mesh closely resembled that data mesh, (as described in Section 6.5.3), then the model resulting from that trial was saved as a JSON file. The data of the 20 models were then converted to a spreadsheet in CSV format for analysis (see Table A3.1) using the `ModellerExporter.py` script (Appendix 2). Table A3.1 shows the computed mean and standard deviation for the 10 trials for each attribute, for each of the two models. The columns labeled ‘Mean EAS’ and ‘STDEV EAS’ are the resulting mean and standard deviations for 10 repetitions of modelling EAS\_M20; the columns ‘Mean EUR’ and ‘STDEV EUR’ are the counterparts for the 10 repetitions for EUR\_M20). The rows are grouped by facial region, starting with the cranium. Note that CRN\_depth was not modified since there was no data provided for cranial depth.

Recall that unsigned attributes vary from 0.0 to 1.0 (a range of 1.0) while signed attributes vary from -1.0 to 1.0 (a range of 2.0). Computing the quotient of one standard deviation and the range for that attribute provides a basis for judging the relative distribution of repeated measurements for each attribute (see the two columns ‘EAS %’ and ‘EUR %’). The majority of the facial attributes had relative standard deviations of less than 10% (and shown with green values). Roughly one third of the attributes showed relative

repeatability of 5% or less. A few other attributes, however, were much less precise, with relative standard deviations greater than 15% of the range (and shown in red in Table A3.1).

*Table A3.1. Repeatability study results for matching the model mesh in the EM to three-dimensional face data (see Section 6.5.3 and Section 7.1).*

attribute	description	region	MEAN EAS	STDEV EAS	%	MEAN EUR	STDEV EUR	%
CRN_depth	cranium depth (Z)	Cranium	0.00	0.00	0.00	0.00	0.00	0.00
CRN_height	cranium height (Y)	Cranium	0.35	0.09	0.05	0.85	0.05	0.03
CRN_width	cranium width at tragon t-t (X)	Cranium	0.31	0.08	0.04	0.62	0.19	0.09
FOR_curvature	forehead curvature (X)	Cranium	-0.49	0.13	0.06	-0.80	0.11	0.05
FOR_protrusion	forehead protrusion (Z)	Cranium	0.14	0.05	0.05	0.84	0.07	0.07
FOR_slope	forehead slope (Z)	Cranium	0.01	0.11	0.06	-0.59	0.16	0.08
FOR_width	forehead width (X)	Cranium	-0.61	0.16	0.08	-0.36	0.17	0.08
SOR_height	supraorbital ridge height (Y)	Cranium	0.65	0.14	0.07	-0.85	0.14	0.07
SOR_protrusion	supraorbital ridge protrusion (Z)	Cranium	0.26	0.05	0.05	0.89	0.03	0.03
TMP_width	temple width (X)	Cranium	0.46	0.09	0.04	0.93	0.05	0.02
ECF_weight	epicanthal fold	Eye	0.70	0.04	0.04	0.00	0.00	0.00
ENC_angle	endocanthal angle	Eye	0.59	0.14	0.07	0.48	0.20	0.10
ENC_x	endocanthal distance (en-en)	Eye	-0.47	0.10	0.05	-0.36	0.14	0.07
ENC_y	endocanthal height (Y)	Eye	-0.62	0.10	0.05	-0.53	0.20	0.10
EXC_x	exocanthal distance (ex-ex)	Eye	0.67	0.07	0.03	0.42	0.16	0.08
EXC_y	exocanthal height (Y)	Eye	0.46	0.09	0.05	0.40	0.15	0.07
IPC_weight	infrapalpebral convexity	Eye	0.19	0.07	0.07	0.28	0.11	0.11
IPD_x	interpupillary distance (X)	Eye	0.41	0.06	0.03	0.67	0.17	0.08
IPS_weight	infrapalpebral sulcus	Eye	0.51	0.17	0.17	0.55	0.10	0.10
PF_inclination	palpebral fissure inclination	Eye	0.40	0.13	0.13	0.12	0.06	0.06
PF_width	palpebral fissure width (Y)	Eye	0.57	0.08	0.04	0.51	0.15	0.07
PMS_weight	palpebromalar sulcus	Eye	0.72	0.15	0.15	0.64	0.18	0.18
PNS_depth	palpebronasal sulcus	Eye	0.68	0.09	0.09	0.53	0.18	0.18
PTF_depth	palpebrotemporal fossa	Eye	0.54	0.17	0.17	0.57	0.19	0.19
SPC_weight	suprapalpebral convexity	Eye	0.43	0.06	0.06	0.36	0.17	0.17
SPS_weight	suprapalpebral sulcus	Eye	0.61	0.12	0.12	0.35	0.08	0.08
STF_weight	supratarsal fold	Eye	0.21	0.08	0.08	0.34	0.12	0.12
CHK_protrusion	cheek protrusion (Z)	Face	0.57	0.05	0.02	0.46	0.13	0.06
CHK_width	cheek width (X)	Face	-0.36	0.11	0.06	-0.62	0.16	0.08
MAX_protrusion	maxillary protrusion (Z)	Face	-0.71	0.17	0.09	-0.42	0.06	0.03
TRG_y	tragus placement (Y)	Face	-0.52	0.09	0.05	0.28	0.18	0.09
TRG_z	tragus placement (Z)	Face	-0.78	0.19	0.10	-0.94	0.08	0.04
ZYG_protrusion	cheek bone protrusion (Z)	Face	-0.15	0.03	0.02	-0.55	0.18	0.09

attribute	description	region	MEAN EAS	STDEV EAS	%	MEAN EUR	STDEV EUR	%
<b>ZYG_width</b>	<b>cheek bone width (X)</b>	<b>Face</b>	<b>0.56</b>	<b>0.08</b>	<b>0.04</b>	<b>0.73</b>	<b>0.14</b>	<b>0.07</b>
GNA_x	gnathion width (X)	Jaw	-0.50	0.20	0.10	-0.44	0.15	0.08
GNA_y	gnathion placement (Y)	Jaw	0.88	0.07	0.04	0.92	0.08	0.04
GNA_z	gnathion placement (Z)	Jaw	-0.20	0.04	0.02	-0.60	0.11	0.06
<b>GON_x</b>	<b>gonion placement (X)</b>	<b>Jaw</b>	<b>0.44</b>	<b>0.17</b>	<b>0.08</b>	<b>0.38</b>	<b>0.08</b>	<b>0.04</b>
GON_y	gonion placement (Y)	Jaw	0.49	0.23	0.11	0.24	0.14	0.07
LMS_weight	labiomental sulcus weight	Jaw	0.11	0.10	0.10	0.30	0.19	0.19
MDB_protrusion	mandibular protrusion (Z)	Jaw	-0.54	0.18	0.09	-0.98	0.04	0.02
MDB_width	mandibular width (X)	Jaw	-0.47	0.10	0.05	-0.63	0.15	0.08
NCK_protrusion	neck protrusion	Jaw	0.55	0.11	0.11	0.29	0.15	0.15
NCK_width	neck breadth	Jaw	0.52	0.09	0.09	0.35	0.11	0.11
<b>CH_x</b>	<b>mouth width (ch-ch)</b>	<b>Mouth</b>	<b>0.59</b>	<b>0.09</b>	<b>0.04</b>	<b>0.67</b>	<b>0.09</b>	<b>0.05</b>
CPB_depth	cupid's bow depth (Z)	Mouth	0.49	0.15	0.08	0.79	0.07	0.03
CPB_width	cupid's bow width (X)	Mouth	-0.45	0.13	0.07	0.25	0.12	0.06
IL_convexity	inf. labium fullness (Z)	Mouth	0.33	0.19	0.19	0.25	0.18	0.18
IL_protrusion	inf. labium protrusion (Z)	Mouth	0.15	0.17	0.08	-0.49	0.22	0.11
IL_thickness	inf. labium thickness (Y)	Mouth	0.53	0.17	0.09	0.61	0.13	0.07
ILT_fullness	inf. labium tubercle fullness	Mouth	-0.50	0.17	0.08	-0.36	0.17	0.09
IVB_curve	inf. vermillion border curvature	Mouth	-0.38	0.10	0.05	-0.39	0.10	0.05
PHL_length	philtrum length (Y)	Mouth	-0.47	0.09	0.04	-0.54	0.11	0.05
PHL_protrusion	philtrum protrusion at sn (Z)	Mouth	0.34	0.19	0.19	0.84	0.10	0.10
PHL_width	philtrum width (X)	Mouth	0.56	0.09	0.05	0.57	0.19	0.09
SL_convexity	sup. labium convexity (Z)	Mouth	0.65	0.14	0.14	0.69	0.15	0.15
SL_protrusion	sup. labium protrusion (Z)	Mouth	-0.32	0.17	0.08	-0.57	0.20	0.10
SL_thickness	sup. labium thickness (Y)	Mouth	-0.43	0.14	0.07	-0.62	0.12	0.06
SLT_fullness	sup. labium tubercle fullness	Mouth	-0.33	0.15	0.07	-0.50	0.18	0.09
SVB_curve	sup.vermillion border curvature	Mouth	-0.35	0.16	0.08	-0.55	0.19	0.10
ALA_contour	alar contour (sulcus)	Nose	0.52	0.12	0.06	0.55	0.11	0.05
ALA_drop	alar drop (Y)	Nose	0.36	0.16	0.16	0.15	0.10	0.10
<b>ALA_width</b>	<b>alar width al-al (X)</b>	<b>Nose</b>	<b>-0.39</b>	<b>0.17</b>	<b>0.08</b>	<b>-0.55</b>	<b>0.17</b>	<b>0.08</b>
COL_drop	columella drop (Y)	Nose	0.24	0.06	0.06	0.24	0.16	0.16
COL_show	columellar show (Y)	Nose	0.70	0.09	0.09	0.77	0.16	0.16
COL_width	columella width (X)	Nose	0.45	0.20	0.10	0.73	0.08	0.04
<b>DSM_length</b>	<b>dorsum length n-sn (Y)</b>	<b>Nose</b>	<b>0.21</b>	<b>0.04</b>	<b>0.02</b>	<b>0.26</b>	<b>0.05</b>	<b>0.03</b>
DSM_protrusion	dorsum protrusion (Z)	Nose	-0.91	0.07	0.03	0.18	0.05	0.03
DSM_width	dorsum width (X)	Nose	0.56	0.14	0.07	0.68	0.08	0.04
RAD_protrusion	radix protrusion (Z)	Nose	-0.92	0.08	0.04	0.59	0.11	0.05
RAD_width	radix width (X)	Nose	0.66	0.12	0.06	0.67	0.12	0.06
SID_slope	sidewall slope (X)	Nose	-0.64	0.19	0.10	-0.49	0.18	0.09
TIP_drop	infratip lobule drop (Y)	Nose	0.14	0.07	0.07	0.19	0.08	0.08
TIP_inclination	nasal tip inclination	Nose	0.41	0.18	0.09	-0.25	0.15	0.07
TIP_pointed	tip pointed	Nose	-0.65	0.27	0.14	-0.50	0.13	0.07
TIP_protrusion	tip protrusion (Z)	Nose	-0.48	0.16	0.08	0.23	0.07	0.04
TIP_width	tip width (X)	Nose	0.46	0.10	0.05	0.57	0.12	0.06

These low-precision attributes were all shape attributes, in fact, such as *PTF\_depth* (the depth of the palpebrotemporal fossa, the slight hollow lateral to the eye), and *IL\_convexity* (the fullness of the inferior labium). From the user perspective, these shape attributes required larger adjustments in order to be noticed.

The five rows shown in bold are for facial attributes that have direct anthropometric counterparts (e.g., *DSM\_length*, corresponds to the conventional measurement *n-sn*). These attributes were then calibrated into dimensions in millimetres, and since those dimensional attributes have reasonable repeatability (precision), they were important in the subsequent accuracy study; see Section 7.1 and Appendix 3.2 for discussion.

### **Appendix 3.2 Accuracy Study**

While most of the attributes in the Ethnicity Modeller have no direct counterpart in conventional anthropometry, five attributes do: *ALA\_width*, *CH\_x*, *DSM\_length*, *GON\_x*, and *ZYG\_width*.

The attribute *ALA\_width* corresponds to conventional measurement *al-al*, *DSM\_length* with *n-sn*, and so forth. Each of these dimensional attributes were calibrated by measuring the basis shapes corresponding to the minimum and maximum extremes of that attribute in Maya (using the measurement tool). This permitting conversion of any attribute value from -1.0 to 1.0 to millimetres by linear interpolation within those extremes for that measurement.

The measurement data provided by the Ethnicity Modeller was compared with the conventional anthropometric measurements in (Farkas et al., 2005). For each of the five attributes the computed mean (over 10 trials) from Table A3.1 for EAS\_M20 was compared with the corresponding physical measurements for the two EAS ethnotypes Singaporean Chinese males (Farkas et al., 2005, table 19a), and Japanese Males (table 22a). Then, to compare the Ethnicity Modeller results for the averaged EUR model (EUR\_M20), two EUR data sets were used from (Farkas et al., 2005): North American

white male (table 1) and German males (table 6a). The comparisons for EAS are shown in Table A3.2 and for EUR in Table A3.3. The mean of 10 trials using the Ethnicity Modeller is shown in terms of the attribute value, its equivalent in millimetres, then compared to the the corresponding values from Farkas et al. (2005).

Consider first the accuracy with which the *EM* measures the EAS face for these five attributes (Table A3.2). The mean value of the attribute *ALA\_width (al-al)* for example was  $-0.39 \pm 0.17$ , and converted to millimetres was  $36.8 \pm 1.6$  mm, and then compared with *al-al* for Singaporean Chinese (CN) and Japanese (JP) data.

Note that the model of EAS\_M20 is based on the composite photogrammetric data of 20 male EAS, but not specifically CN or JP males. Nonetheless, the EAS\_M20 model data generally matches the anthropometric data very well, i.e., to within one standard deviation for nose width *ALA\_width* (which corresponds to *al-al*), mouth width *CH\_x* (corresponding to *ch-ch*), and nose length *DSM\_length* (corresponding to *n-sn*). The matching is similarly close for the jaw width *GON\_x* (corresponding to *go-go*) and zygomatic width *ZYG\_x* (corresponding to *zy-zy*) for CH, but show an underestimation of nearly 13 mm for JP. These discrepancies are shown in red in Table A3.2. This will be discussed after showing the corresponding data for the EUR\_M20 measurements, since the trends are very similar.

*Table A3.2. Accuracy study of five anthropometric attributes, comparing measurements taken with the Ethnicity Modeller for EAS\_M20 with the corresponding measurements from Farkas et al. (2005) for Singaporean Chinese (CN) and Japanese (JP) males.*

Attribute	EAS_M20 (value)	EAS_M20 (mm)	Farkas CN (mm)	difference (mm)	Farkas JP (mm)	difference (mm)
<b><i>ALA_width (al-al)</i></b>	$-0.39 \pm 0.17$	$36.8 \pm 1.6$	$39.2 \pm 2.9$	-2.4	$38.2 \pm 2.5$	-1.4
<b><i>CH_x (ch-ch)</i></b>	$0.59 \pm 0.09$	$51.1 \pm 0.6$	$49.6 \pm 2.8$	1.5	$48.4 \pm 3.5$	2.7
<b><i>DSM_length (n-sn)</i></b>	$0.21 \pm 0.04$	$56.3 \pm 0.5$	$53.8 \pm 3.0$	2.5	$56.9 \pm 4.9$	-0.6
<b><i>GON_x (go-go)</i></b>	$0.44 \pm 0.17$	$104.6 \pm 1.0$	$107.3 \pm 5.6$	-2.7	$117.3 \pm 7.9$	-12.7
<b><i>ZYG_x (zy-zy)</i></b>	$0.56 \pm 0.08$	$139.8 \pm 0.4$	$144.6 \pm 5.6$	-4.8	$147.2 \pm 5.6$	-7.4

Next, the comparison of the EUR\_M20 model with two EUR ethnotypes North American (NA) males and German (DE) males is given in Table A3.3. The difference between the model of EUR\_M20 and the mean anthropometric data for both is less than one standard deviation for *ALA\_width*, *CH\_x*, and *DSM\_width*. Again, there are large differences for the *GON\_x* and *ZYG\_x* attributes (indicated in red). There is insufficient data here to understand in much detail why the *GON\_x* and *ZYG\_x* measurements differ from those in Farkas et al. (2005). What can be offered, however, is that while the literature has very well-defined superficial landmarks such as (the nasal ala *al*, cheilion *ch*, nasion *n*, and subnasale *sn*), the gonion *go* and zygon *zy* are in fact based on the underlying osteology, and there are large disparities in where those superficial points are indicated diagrammatically in Farkas et al. (2005) versus the radiographically-based definition in Swennen (2005). Given that the mandible widens from the chin back towards the ears, any differences in convention for locating the gonion could well underlie a disparity of 6 or 7 millimetres (total, or half that on each side). Similarly, the zygon *zy* is defined osteologically and yet measured on the surface of the face, on the side of the cheek with differing specific locations according to different studies. There is also the potential that the disparities between the *Ethnicity Modeller* values for *go-go* and *zy-zy* reflect actual ethnic differences between EAS\_M20 (the average of 20 EAS provided by stereo-photogrammetry) and the data from Farkas et al. (2005). Likewise, the average comprising EUR\_M20 may reflect a different population than the NA or DE data from (Farkas et al., 2005). Whatever the cause for the disparities shown in red in tables A3.2

*Table A3.3. Corresponding accuracy study for EUR\_M20 compared to the anthropometric measurements from Farkas et al. (2005) for North American male (NA) and German male (DE).*

Attribute	EUR_M20 (value)	EUR_M20 (mm)	Farkas NA (mm)	difference (mm)	Farkas DE (mm)	difference (mm)
<b><i>ALA_width (al-al)</i></b>	-0.55 ± 0.17	35.3 ± 1.6	34.7 ± 2.6	0.6	34.0 ± 2.2	1.3
<b><i>CH_x (ch-ch)</i></b>	0.67 ± 0.09	51.7 ± 0.6	53.3 ± 3.3	-1.6	50.9 ± 3.4	0.8
<b><i>DSM_length (n-sn)</i></b>	0.26 ± 0.05	57.0 ± 0.7	53.3 ± 3.5	4.0	52.0 ± 5.6	5.0
<b><i>GON_x (go-go)</i></b>	0.38 ± 0.04	104.7 ± 0.6	97.1 ± 5.8	7.2	97.6 ± 6.0	6.7
<b><i>ZYG_x (zy-zy)</i></b>	0.73 ± 0.14	140.0 ± 0.4	137.1 ± 4.3	3.6	133.2 ± 7.5	7.5

and A3.3, the differences are of questionable significance as they differ by less than one standard deviation (except for the *go-go* measurement for JP, as noted).

### Appendix 3.3 Quantifying Ethnic Differences

The Ethnicity Modeller was then used to examine the differences between the data sets for the ethnicities EAS and EUR. Each of the 77 attributes were compared pairwise by computing the difference and error propagation since many of these measurements have considerable variance across trials (see Table 3.4). The table was then sorted according to decreasing value of the absolute value of the difference.

*Table A3.4. Computation of the differences in attribute values for two ethnicities. The table is sorted by the absolute value of the difference. Attributes in bold are calibrated and correspond to anthropometry measurements. Attributes of the nose region are highlighted (see text).*

attribute	description	region	EAS	EUR	DIF
<b>RAD_protrusion</b>	radix protrusion (Z)	Nose	-0.92 ± 0.08	0.59 ± 0.11	1.51 ± 0.19
SOR_height	supraorbital ridge height (Y)	Cranium	0.65 ± 0.14	-0.85 ± 0.14	-1.50 ± 0.27
<b>DSM_protrusion</b>	dorsum protrusion (Z)	Nose	-0.91 ± 0.07	0.18 ± 0.05	1.10 ± 0.12
TRG_y	tragus placement (Y)	Face	-0.52 ± 0.09	0.28 ± 0.18	0.80 ± 0.27
<b>TIP_protrusion</b>	tip protrusion (Z)	Nose	-0.48 ± 0.16	0.23 ± 0.07	0.71 ± 0.24
ECF_weight	epicanthal fold	Eye	0.70 ± 0.04	0.00 ± 0.00	-0.70 ± 0.04
FOR_protrusion	forehead protrusion (Z)	Cranium	0.14 ± 0.05	0.84 ± 0.07	0.70 ± 0.12
CPB_width	cupid's bow width (X)	Mouth	-0.45 ± 0.13	0.25 ± 0.12	0.70 ± 0.25
<b>TIP_inclination</b>	nasal tip inclination	Nose	0.41 ± 0.18	-0.25 ± 0.15	-0.66 ± 0.33
IL_protrusion	inf. labium protrusion (Z)	Mouth	0.15 ± 0.17	-0.49 ± 0.22	-0.63 ± 0.38
SOR_protrusion	supraorbital ridge protrusion (Z)	Cranium	0.26 ± 0.05	0.89 ± 0.03	0.63 ± 0.08
FOR_slope	forehead slope (Z)	Cranium	0.01 ± 0.11	-0.59 ± 0.16	-0.60 ± 0.27
PHL_protrusion	philtrum protrusion at sn (Z)	Mouth	0.34 ± 0.19	0.84 ± 0.10	0.50 ± 0.30
CRN_height	cranium height (Y)	Cranium	0.35 ± 0.09	0.85 ± 0.05	0.49 ± 0.14
TMP_width	temple width (X)	Cranium	0.46 ± 0.09	0.93 ± 0.05	0.46 ± 0.13
MDB_protrusion	mandibular protrusion (Z)	Jaw	-0.54 ± 0.18	-0.98 ± 0.04	-0.44 ± 0.22
ZYG_protrusion	cheek bone protrusion (Z)	Face	-0.15 ± 0.03	-0.55 ± 0.18	-0.40 ± 0.21
GNA_z	gnathion placement (Z)	Jaw	-0.20 ± 0.04	-0.60 ± 0.11	-0.39 ± 0.15
CRN_width	cranium width at trignon t-t (X)	Cranium	0.31 ± 0.08	0.62 ± 0.19	0.32 ± 0.27
FOR_curvature	forehead curvature (X)	Cranium	-0.49 ± 0.13	-0.80 ± 0.11	-0.31 ± 0.24
CPB_depth	cupid's bow depth (Z)	Mouth	0.49 ± 0.15	0.79 ± 0.07	0.30 ± 0.22
MAX_protrusion	maxillary protrusion (Z)	Face	-0.71 ± 0.17	-0.42 ± 0.06	0.29 ± 0.24
PF_inclination	palpebral fissure inclination	Eye	0.40 ± 0.13	0.12 ± 0.06	-0.29 ± 0.19

attribute	description	region	EAS	EUR	DIF
COL_width	columella width (X)	Nose	0.45 ± 0.20	0.73 ± 0.08	0.28 ± 0.29
CHK_width	cheek width (X)	Face	-0.36 ± 0.11	-0.62 ± 0.16	-0.27 ± 0.27
IPD_x	interpupillary distance (X)	Eye	0.41 ± 0.06	0.67 ± 0.17	0.26 ± 0.23
NCK_protrusion	neck protrusion	Jaw	0.55 ± 0.11	0.29 ± 0.15	-0.26 ± 0.25
SPS_weight	suprapalpebral sulcus	Eye	0.61 ± 0.12	0.35 ± 0.08	-0.26 ± 0.20
GON_y	gonion placement (Y)	Jaw	0.49 ± 0.23	0.24 ± 0.14	-0.25 ± 0.37
FOR_width	forehead width (X)	Cranium	-0.61 ± 0.16	-0.36 ± 0.17	0.25 ± 0.33
SL_protrusion	sup. labium protrusion (Z)	Mouth	-0.32 ± 0.17	-0.57 ± 0.20	-0.25 ± 0.37
EXC_x	exocanthal distance (ex-ex)	Eye	0.67 ± 0.07	0.42 ± 0.16	-0.24 ± 0.23
ALA_drop	alar drop (Y)	Nose	0.36 ± 0.16	0.15 ± 0.10	-0.21 ± 0.25
SVB_curve	sup. vermillion border curvature	Mouth	-0.35 ± 0.16	-0.55 ± 0.19	-0.20 ± 0.35
LMS_weight	labiomenal sulcus weight	Jaw	0.11 ± 0.10	0.30 ± 0.19	0.19 ± 0.29
SL_thickness	sup. labium thickness (Y)	Mouth	-0.43 ± 0.14	-0.62 ± 0.12	-0.19 ± 0.26
SLT_fullness	sup. labium tubercle fullness	Mouth	-0.33 ± 0.15	-0.50 ± 0.18	-0.17 ± 0.33
NCK_width	neck breadth	Jaw	0.52 ± 0.09	0.35 ± 0.11	-0.17 ± 0.20
<b>ZYG_width</b>	<b>cheek bone width (X)</b>	<b>Face</b>	<b>0.56 ± 0.08</b>	<b>0.73 ± 0.14</b>	<b>0.17 ± 0.21</b>
<b>ALA_width</b>	<b>alar width al-al (X)</b>	<b>Nose</b>	<b>-0.39 ± 0.17</b>	<b>-0.55 ± 0.17</b>	<b>-0.16 ± 0.33</b>
MDB_width	mandibular width (X)	Jaw	-0.47 ± 0.10	-0.63 ± 0.15	-0.16 ± 0.25
TRG_z	tragus placement (Z)	Face	-0.78 ± 0.19	-0.94 ± 0.08	-0.15 ± 0.27
TIP_pointed	tip pointed	Nose	-0.65 ± 0.27	-0.50 ± 0.13	0.15 ± 0.41
SID_slope	sidewall slope (X)	Nose	-0.64 ± 0.19	-0.49 ± 0.18	0.15 ± 0.37
PNS_depth	palpebronasal sulcus	Eye	0.68 ± 0.09	0.53 ± 0.18	-0.15 ± 0.27
ILT_fullness	inf. labium tubercle fullness	Mouth	-0.50 ± 0.17	-0.36 ± 0.17	0.14 ± 0.34
STF_weight	supratarsal fold	Eye	0.21 ± 0.08	0.34 ± 0.12	0.13 ± 0.20
DSM_width	dorsum width (X)	Nose	0.56 ± 0.14	0.68 ± 0.08	0.12 ± 0.22
ENC_angle	endocanthal angle	Eye	0.59 ± 0.14	0.48 ± 0.20	-0.11 ± 0.34
CHK_protrusion	cheek protrusion (Z)	Face	0.57 ± 0.05	0.46 ± 0.13	-0.11 ± 0.17
ENC_x	endocanthal distance (en-en)	Eye	-0.47 ± 0.10	-0.36 ± 0.14	0.11 ± 0.23
TIP_width	tip width (X)	Nose	0.46 ± 0.10	0.57 ± 0.12	0.11 ± 0.22
ENC_y	endocanthal height (Y)	Eye	-0.62 ± 0.10	-0.53 ± 0.20	0.09 ± 0.29
IPC_weight	infrapalpebral convexity	Eye	0.19 ± 0.07	0.28 ± 0.11	0.09 ± 0.18
<b>CH_x</b>	<b>mouth width (ch-ch)</b>	<b>Mouth</b>	<b>0.59 ± 0.09</b>	<b>0.67 ± 0.09</b>	<b>0.08 ± 0.18</b>
IL_convexity	inf. labium fullness (Z)	Mouth	0.33 ± 0.19	0.25 ± 0.18	-0.08 ± 0.37
IL_thickness	inf. labium thickness (Y)	Mouth	0.53 ± 0.17	0.61 ± 0.13	0.08 ± 0.31
SPC_weight	suprapalpebral convexity	Eye	0.43 ± 0.06	0.36 ± 0.17	-0.08 ± 0.23
PMS_weight	palpebromalar sulcus	Eye	0.72 ± 0.15	0.64 ± 0.18	-0.08 ± 0.34
PHL_length	philtrum length (Y)	Mouth	-0.47 ± 0.09	-0.54 ± 0.11	-0.07 ± 0.20
GNA_x	gnathion width (X)	Jaw	-0.50 ± 0.20	-0.44 ± 0.15	0.06 ± 0.35
PF_width	palpebral fissure width (Y)	Eye	0.57 ± 0.08	0.51 ± 0.15	-0.06 ± 0.23
<b>GON_x</b>	<b>gonion placement (X)</b>	<b>Jaw</b>	<b>0.44 ± 0.17</b>	<b>0.38 ± 0.08</b>	<b>-0.06 ± 0.25</b>
COL_show	columellar show (Y)	Nose	0.70 ± 0.09	0.77 ± 0.16	0.06 ± 0.25
EXC_y	exocanthal height (Y)	Eye	0.46 ± 0.09	0.40 ± 0.15	-0.06 ± 0.24
TIP_drop	infratip lobule drop (Y)	Nose	0.14 ± 0.07	0.19 ± 0.08	0.05 ± 0.15
<b>DSM_length</b>	<b>dorsum length n-sn (Y)</b>	<b>Nose</b>	<b>0.21 ± 0.04</b>	<b>0.26 ± 0.05</b>	<b>0.04 ± 0.09</b>
IPS_weight	infrapalpebral sulcus	Eye	0.51 ± 0.17	0.55 ± 0.10	0.04 ± 0.27



attribute	description	region	EAS	EUR	DIF
GNA_y	gnathion placement (Y)	Jaw	0.88 ± 0.07	0.92 ± 0.08	0.04 ± 0.15
SL_convexity	sup. labium convexity (Z)	Mouth	0.65 ± 0.14	0.69 ± 0.15	0.04 ± 0.30
PTF_depth	palpebrotemporal fossa	Eye	0.54 ± 0.17	0.57 ± 0.19	0.03 ± 0.36
ALA_contour	alar contour (sulcus)	Nose	0.52 ± 0.12	0.55 ± 0.11	0.03 ± 0.23
IVB_curve	inf. vermillion border curvature	Mouth	-0.38 ± 0.10	-0.39 ± 0.10	-0.02 ± 0.20
PHL_width	philtrum width (X)	Mouth	0.56 ± 0.09	0.57 ± 0.19	0.01 ± 0.28
RAD_width	radix width (X)	Nose	0.66 ± 0.12	0.67 ± 0.12	0.01 ± 0.25
COL_drop	columella drop (Y)	Nose	0.24 ± 0.06	0.24 ± 0.16	0.00 ± 0.22
CRN_depth	cranium depth (Z)	Cranium	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00

The attribute *RAD\_protrusion* (the protrusion of the radix, or base, of the nose in Z) shows the greatest difference, from  $-0.92 \pm 0.08$  for EAS versus  $0.59 \pm 0.11$  for EUR. It should be kept in mind that this is a difference in attribute value within the -1.0 to 1.0 range, and not an absolute difference in millimetres of the protrusion of the radix, for the two ethnotypes. But consider the rows highlighted in yellow in Table A3.4. These first four describe the characteristics of the EUR versus EAS nose in profile: the EUR nose has greater protrusion at the radix, dorsum, and tip, and the tip is less inclined. These all were ranked higher in this table than the two attributes *ALA\_width* and *DSM\_length*, the familiar attributes defining the nasal index which is usually used to differentiate EAS and EUR (Section 2.2).

Note that *SOR\_height* (the placement of the supraorbital ridge in Y above the eyes) shows a larger numerical difference between EAS and EUR than does *ECF\_weight* (the epicanthal fold). But it does not follow, however, that the epicanthal fold is a less salient feature for distinguishing EAS from EUR, of course (see Section 7.2.1 for discussion).

### Appendix A3.4 Measuring Individuals

Stereo-photogrammetric mesh data for 10 individual EAS males, and 10 individual EUR males were separately modelled. In a follow-on to the accuracy study in Appendix A3.2, the averaged attribute values for these two sample populations were compared with the anthropometric data provided in (Farkas et al., 2005). Table A3.5 shows the computed mean attribute value, the corresponding dimension (in millimetres), and the physically-

measured anthropometric means for 20 Singaporean Chinese (CN) and 20 Japanese. The *Ethnicity Modeller* estimate of alar width *ALA\_width* (or al-al) of  $38.8 \pm 2.9$  mm was remarkably close to the physical anthropometric measurements (differing from CN by -0.4 mm and from JP by 0.6 mm). Only in comparing gonion width *GON\_x* (or go-go) for the JP data was there a sizeable disparity (shown in red), which was also noted for in Table 3.2 for JP.

*Table A3.5. Comparison of mean of 10 EAS male individuals with the corresponding measurements from Farkas et al. (2005) for Singaporean Chinese (CN) and Japanese (JP) males.*

Attribute	10 EAS	10 EAS (mm)	Farkas CN (mm)	difference (mm)	Farkas JP (mm)	difference (mm)
<b><i>ALA_width (al-al)</i></b>	$-0.18 \pm 0.16$	$38.8 \pm 1.5$	$39.2 \pm 2.9$	-0.4	$38.2 \pm 2.5$	0.6
<b><i>CH_x (ch-ch)</i></b>	$0.57 \pm 0.31$	$51.0 \pm 2.2$	$49.6 \pm 2.8$	1.4	$48.4 \pm 3.5$	2.6
<b><i>DSM_length (n-sn)</i></b>	$0.20 \pm 0.26$	$56.2 \pm 3.5$	$53.8 \pm 3.0$	2.4	$56.9 \pm 4.9$	-0.7
<b><i>GON_x (go-go)</i></b>	$0.13 \pm 0.41$	$102.8 \pm 2.5$	$107.3 \pm 5.6$	-4.5	$117.3 \pm 7.9$	-14.5
<b><i>ZYG_x (zy-zy)</i></b>	$0.86 \pm 0.27$	$141.3 \pm 1.4$	$144.6 \pm 5.6$	-3.3	$147.2 \pm 5.6$	-5.9

The corresponding comparisons between the 10 EUR measured by the EM and the physical measurements in (Farkas et al., 2005) for 20 North American (NA) males and 20 German (DE) males are shown in Table A3.6. The *EM* measurements are even closer than for the EAS.

*Table A3.6. Comparison of mean of 10 EUR male individuals with the corresponding measurements North American male (NA) and German male (DE) in (Farkas et al.,*

Attribute	10 EUR (value)	10 EUR (mm)	Farkas NA (mm)	difference (mm)	Farkas DE (mm)	difference (mm)
<b><i>ALA_width (al-al)</i></b>	$-0.68 \pm 0.26$	$34.0 \pm 2.47$	$34.7 \pm 2.6$	-0.7	$34.0 \pm 2.2$	0.0
<b><i>CH_x (ch-ch)</i></b>	$0.24 \pm 0.32$	$48.7 \pm 2.2$	$53.3 \pm 3.3$	-4.6	$50.9 \pm 3.4$	-2.2
<b><i>DSM_length (n-sn)</i></b>	$0.05 \pm 0.29$	$54.2 \pm 3.9$	$53.3 \pm 3.5$	1.2	$52.0 \pm 5.6$	2.2
<b><i>GON_x (go-go)</i></b>	$-0.11 \pm 0.36$	$101.3 \pm 2.2$	$97.1 \pm 5.8$	4.2	$97.6 \pm 6.0$	3.7
<b><i>ZYG_x (zy-zy)</i></b>	$0.38 \pm 0.47$	$138.9 \pm 2.4$	$137.1 \pm 4.3$	1.8	$133.2 \pm 7.5$	5.7